11. Point and Interval Estimation

Learning Unit	Learning Objective
Statistics Area	
Point and Interval Estimation	
21. Sampling distribution and point estimates	21.1 recognise the concepts of sample statistics and population parameters 21.2 recognise the sampling distribution of the sample mean from a random sample of size n 21.3 recognise the concept of point estimates including the sample mean, sample variance and sample proportion 21.4 recognise Central Limit Theorem
22. Confidence interval for a population mean	22.1 recognise the concept of confidence interval 22.2 find the confidence interval for a population mean
23. Confidence interval for a population proportion	23.1 find an approximate confidence interval for a population proportion

Summary

A. About population mean

For a normal population, i.e. $X \sim N(\mu, \sigma^2)$, based on a sample size n, $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$. It should be noted that the results are true for samples of any sizes.

For a non-normal population with the population mean μ and a known variance σ^2, if the sample size is large ($n \geq 30$), Central Limit Theorem can be used. $\bar{X}$ is approximately normal and $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

For a normal or non-normal population with the population mean μ and <u>*unknown*</u> variance σ^2, if the sample size is large ($n \geq 30$), X is approximately normal and $\bar{X} \sim N\left(\mu, \frac{s^2}{n}\right)$ where $\frac{s}{\sqrt{n}}$ is called the standard error of the sample and is denoted by $SE(\bar{x})$. (For normal population and small sample size, t-distribution is used.)

Conditions	95% confidence interval for μ	99% confidence interval for μ
Normal population · with known variance σ^2 · large or small sample size n · sample mean $\bar{x}$	$(\bar{x} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96\frac{\sigma}{\sqrt{n}})$	$(\bar{x} - 2.575\frac{\sigma}{\sqrt{n}}, \bar{x} + 2.575\frac{\sigma}{\sqrt{n}})$
Non-normal population · with known variance σ^2 · large sample size n ($n \geq 30$) · sample mean $\bar{x}$	$(\bar{x} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96\frac{\sigma}{\sqrt{n}})$	$(\bar{x} - 2.575\frac{\sigma}{\sqrt{n}}, \bar{x} + 2.575\frac{\sigma}{\sqrt{n}})$
Non-normal population · with unknown variance σ^2 · large sample size n ($n \geq 30$) · sample mean $\bar{x}$ · sample variance s^2	$(\bar{x} - 1.96\frac{s}{\sqrt{n}}, \bar{x} + 1.96\frac{s}{\sqrt{n}})$	$(\bar{x} - 2.575\frac{s}{\sqrt{n}}, \bar{x} + 2.575\frac{s}{\sqrt{n}})$

Provided by dse.life

B. About population proportion

A number of random samples, each of size n, are drawn from a parent population. If the proportion of successes for each sample is $p_s = \frac{x}{n}$, these proportions form a distribution called the sampling distribution of proportions P_s (P_s denotes a distribution). When n is sufficiently large, the distribution of P_s is approximately normal and $P_s \sim N\left(p, \frac{p(1-p)}{n}\right)$ where $\frac{p(1-p)}{n}$ is the standard error of the proportions. The larger the sample size n is, the better is the approximation. Since p is not known, we use p_s to approximate p.

The approximate confidence interval for the population proportion p:

	Conditions
	large sample size n and sample proportion p_s
95% confidence interval	$\left(p_s - 1.96\sqrt{\frac{p_s(1-p_s)}{n}}, p_s + 1.96\sqrt{\frac{p_s(1-p_s)}{n}}\right)$
99% confidence interval	$\left(p_s - 2.575\sqrt{\frac{p_s(1-p_s)}{n}}, p_s + 2.575\sqrt{\frac{p_s(1-p_s)}{n}}\right)$

The sample size, n, should satisfies $\frac{1.96}{\sqrt{n}} \le width$ for 95% C.I. and $\frac{2.575}{\sqrt{n}} \le width$ for 99% C.I..

Section A

1. In an estate, Peter wants to study the proportion p of households who keep pets. He conducts a survey of a random sample of 64 households and finds that an approximate $\beta\%$ confidence interval for p is $(0.0915, 0.3085)$.
   (a) Find
      (i) the sample proportion of households who keep pets,
      (ii) β.
   (b) Using the sample proportion obtained in (a)(i), find the least number of households such that the probability of at least 1 of these households who keeps pets is greater than 0.999.

   (6 marks) (2018 DSE-MATH-M1 Q2)

2. There are many packs of seeds and each pack contains 100 seeds. Let p be the population proportion of seeds that germinate in a pack.
   (a) A pack of seeds is randomly selected, 64 seeds germinate. Find an approximate 95% confidence interval for p.
   (b) It is given that the proportion of seeds that germinate in these packs of seeds follows a normal distribution with a mean of p and a standard deviation of 0.05. Find the least sample size to be taken such that the width of a 90% confidence interval for p is less than 0.04.

   (7 marks) (2016 DSE-MATH-M1 Q4)

3. The manager of a fitness centre wants to promote aerobic classes.
   (a) The manager randomly selected 200 Hong Kong residents and found out that 80 of them had taken part in aerobic classes. Let p be the proportion of Hong Kong residents who had taken part in aerobic classes. Find an approximate 95% confidence interval for p.
   (b) The manager wants to randomly select n Hong Kong residents and invite them to take part in a free aerobic class. The probability that an invited resident will show up is 0.85. Let X be the proportion of the n invited residents who will show up. Assume that X can be modelled by a normal distribution with mean 0.85 and variance $\frac{0.85(1-0.85)}{n}$. Find the maximum number of n such that the probability that more than 100 invited residents will show up is less than 0.05.

   (7 marks) (2014 DSE-MATH-M1 Q9)

Provided by dse.life

4. In a random sample of 120 swimmers in a certain beach, 75 of them are not satisfied with the water quality of the beach. Let p be the population proportion of the swimmers in this beach who are not satisfied with the water quality of the beach. Find an approximate 90% confidence interval for p.

(4 marks) (2013 DSE-MATH-M1 Q6)

5. The lifetime of a randomly selected LED bulb produced by a manufacturer is assumed to be normally distributed with mean μ hours and standard deviation 5000 hours. It is known that 96.41% of the bulbs will have a lifetime shorter than 39000 hours.
   (a) Find the value of μ.
   (b) Suppose a random sample of 100 bulbs is drawn. Find the probability that the mean lifetime of the sample lies between 30200 hours and 30800 hours.
   (c) The manufacturer wants to select another random sample of n bulbs such that the probability that the mean lifetime of the sample exceeding 28500 hours is at least 0.985. Find the least value of n.

(7 marks) (2013 DSE-MATH-M1 Q9)

6. The weights (in kg) of the students in a school can be modelled by the normal distribution with mean 67 and standard deviation 15. A random sample of 36 students is taken.
   (a) Find the probability that the mean weight of the 36 students is over 70 kg.
   (b) It is found that 9 students in the sample like French fries. Find an approximate 95% confidence interval for the proportion of students in the school who like French fries.

(5 marks) (2012 DSE-MATH-M1 Q6)

7. A random sample of size 10 is drawn from a normal population with mean μ and variance 8. Let $\overline{X}$ be the mean of the sample.
   (a) Calculate $\operatorname{Var}(2\overline{X}+7)$.
   (b) Suppose the mean of the sample is 50. Construct a 97% confidence interval for μ.

(5 marks) (PP DSE-MATH-M1 Q6)

8. A political party studied the public view on a certain government policy. A random sample of 150 people was taken and 57 of them supported this policy
   (a) Estimate the population proportion supporting this policy.
   (b) Find an approximate 90% confidence interval for the population proportion.

(4 marks) (SAMPLE DSE-MATH-M1 Q3)

9. A manufacturer produces a large batch of light bulbs, with a mean lifetime of 640 hours and a standard deviation of 40 hours. A random sample of 25 bulbs is taken. Find the probability that the sample mean lifetime of the 25 bulbs is greater than 630 hours.

(5 marks) (SAMPLE DSE-MATH-M1 Q5)

Section B

10. The daily times spent on homework of the students in a school follow a normal distribution with a mean of μ hours and a standard deviation of 0.4 hours.
    (a) A survey is conducted in the school to estimate μ.
        (i) A sample of 40 students in the school is randomly selected and their daily times spent on homework are recorded below:

Daily time sent (x hours)	Number of student
$0.5 < x \leq 1.0$	11
$1.0 < x \leq 1.5$	13
$1.5 < x \leq 2.0$	8
$2.0 < x \leq 2.5$	5
$2.5 < x \leq 3.0$	3

        Find a 90 % confidence interval for μ.
        (ii) Find the least sample size to be taken such that the width of a 97 % confidence interval for μ is at most 0.3.

(7 marks)

    (b) Suppose that $\mu = 1.48$. If the daily time spent on homework of a student exceeds 2 hours, then the student has to attend homework guidance class.
        (i) If a student is randomly selected from the school, find the probability that the student has to attend homework guidance class.
        (ii) A sample of 15 students is now randomly drawn from the school and their daily times spent on homework are examined one by one. Given that more than 1 student in the sample have to attend homework guidance class, find the probability that the 10th student is the 2nd student who has to attend homework guidance class.

(6 marks)

(2017 DSE-MATH-M1 Q9)

Provided by dse.life

11. The speeds of cars passing a checkpoint on a highway follow a normal distribution with a mean of μ km/h and a standard deviation of 16 km/h .

(a) A survey on the speeds of cars to estimate μ is conducted.

(i) A random sample of 25 cars is taken and the stem-and-leaf diagram below shows the distribution of their speeds (in km/h) :

Stem (tens)	Leaf (units)
6	0 0 1 1 1 2 2 3 4 4 5 5 6 6 7
7	1 1 2 3 5 5 6
8	3 6 7

Find a 95% confidence interval for μ .

(ii) Find the least sample size to be taken such that the width of a 97.5% confidence interval for μ is less than 9 .

(7 marks)

(b) Suppose that $\mu = 66$. If the speed of a car passing the checkpoint exceeds 90 km/h , a penalty ticket will be issued.

(i) If a car passes the checkpoint, find the probability that a penalty ticket will be issued.

(ii) If 12 cars pass the checkpoint, find the probability that more than 2 penalty tickets will be issued.

(5 marks)

(2015 DSE-MATH-M1 Q9)

12. The delivery time X (in minutes) of an order received by a pizza restaurant follows a normal distribution with mean μ and standard deviation σ . It is known that 27.43% of the delivery times are longer than 25 minutes and 51.60% of the delivery times fall within 3.5 minutes of μ .

(a) Find μ and σ .

(4 marks)

(b) If the delivery time of an order is longer than k minutes, then a coupon will be given as a compensation to the customer who has made the order. Suppose that a total of 200 orders are received in a day. Assuming independence among delivery times of different orders, find the minimum integral value of k such that the expected number of coupons given out is at most 5 in that day.

(3 marks)

(c) The employees of the pizza restaurant recently received training to improve their efficiency. After training, the delivery time Y (in minutes) of an order follows a normal distribution with mean θ and standard deviation 4.7 .

(i) Manager A draws a random sample of 12 orders and the delivery times (in minutes) are recorded as follows:

22	15	18	21	22	31
20	16	21	19	23	24

Construct a 90% confidence interval for θ .

(ii) Manager B is going to draw another random sample of n orders. He requires that the probability that the mean delivery time of the n orders falls within 3 minutes of θ be greater than 0.99. Find the minimum value of n to meet his requirement.

(6 marks)

(2014 DSE-MATH-M1 Q12)

Provided by dse.life

13. The cholesterol levels (in suitable units) of the adults in a city are assumed to be normally distributed with mean μ and variance σ^2. From a random sample of 49 adults, a 95% confidence interval for μ is found to be $(4.596, 5.044)$.

(a) (i) Find the value of σ.

(ii) Find the mean of the sample.

(3 marks)

(b) Another sample of 15 adults is randomly selected and their cholesterol levels are recorded as follows:

3.6	3.8	3.9	4.3	4.3	4.5	4.8	5.0
5.1	5.2	5.3	5.5	5.8	6.0	6.4	

The two samples are then combined. Construct a 99% confidence interval for μ using the combined sample.

(4 marks)

(c) A health organisation classifies the cholesterol level of an adult to be low, medium and high if his/her cholesterol value is respectively at most 5.2, between 5.2 and 6.2, and at least 6.2. Suppose $\mu = 4.8$:

(i) Find the probability that the cholesterol level of a randomly selected adult in the city is low.

(ii) A sample of 20 adults is randomly selected in the city. Find the probability that there are more than 17 adults with low cholesterol level and at least 1 adult with medium cholesterol level in this sample.

(5 marks)

(2013 DSE-MATH-M1 Q12)

14. A company provides cable-car service for tourists. Tourists complain that the waiting time for the cable-car is too long. From past experience, the waiting time (in minutes) of a randomly selected tourist follows a normal distribution with mean μ and standard deviation 9 .

(a) The customer service manager of the company conducts a survey on the waiting time to estimate μ .

(i) A random sample of 16 tourists is taken and their waiting times are recorded as below:

56	36	48	63	57	41	50	43
56	55	62	46	55	69	38	50

Construct a 90% confidence interval for μ.

(ii) Find the least sample size to be taken such that the width of the 90% confidence interval for μ is less than 6 minutes.

(7 marks)

(b) Suppose that $\mu = 51.5$. The customer service manager of the company interviews tourists and will give a coupon to a tourist whose waiting time is more than 65 minutes.

(i) Find the probability that he gives less than 2 coupons to the first 10 tourists interviewed.

(ii) Find the probability that the 5th coupon is given to the 20th tourist interviewed.

(6 marks)

(2012 DSE-MATH-M1 Q12)

15. A staff of a school studies the school sick room utilization. The number of visits to the sick room per day on 100 randomly selected school days are recorded as follows:

Number of visits per day	0	1	2	3	4	5	6	7
Frequency	6	12	18	21	20	12	7	4

(a) Find an unbiased estimate of the mean number of visits per day.

(1 mark)

(b) (i) Find the sample proportion of school days with less than 4 visits per day.

(ii) Construct an approximate 95% confidence interval for the proportion of school days with less than 4 visits per day.

(3 marks)

(c) Suppose the number of visits per day follows a Poisson distribution with mean λ. Assume that the unbiased estimate obtained in (a) is used for λ. The sick room is said to be *crowded* on a particular day if there are more than 3 visits on that day.

(i) Find the probability that the sick room is *crowded* on a particular day.

(ii) In a certain week of 5 school days, given that the sick room is *crowded* on at least 2 days, find the probability that the sick room is *crowded* on alternate days in the week.

(6 marks)

(PP DSE-MATH-M1 Q12)

Provided by dse.life

16. The Body Mass Index (BMI) value (in kg/m^2) of children aged 12 in a city are assumed to follow a normal distribution with mean μ kg/m^2 and standard deviation 4.5 kg/m^2.

(a) A random sample of nine children aged 12 is drawn and their BMI values (in kg/m^2) are recorded as follows:

16.0, 18.3, 15.2, 17.8, 19.5, 15.9, 18.6, 22.5, 23.6

(i) Find an unbiased estimate for μ.

(ii) Construct a 95% confidence interval for μ.

(3 marks)

(b) Assume $\mu = 18.7$. If a random sample of 25 children aged 12 is drawn and their BMI values are recorded, find the probability that the sample mean is less than 17.8 kg/m^2.

(4 marks)

(c) A child aged 12 having a BMI value greater than 25 kg/m^2 is said to be *overweight*. Children aged 12 are randomly selected one after another and their BMI values are recorded until two *overweight* children are found. Assume that $\mu = 18.7$.

(i) Find the probability that a selected child is *overweight*.

(ii) Find the probability that more than eight children have to be selected in this sampling process.

(iii) Given that more than eight children will be selected in this sampling process, find the probability that exactly ten children are selected.

(8 marks)

(SAMPLE DSE-MATH-M1 Q14)

2021 DSE Q4

Mary conducts a survey to estimate the proportion p of children in a city who learn recorder. In a random sample of 40 children from the city, 28 of them learn recorder.

(a) (i) Find the sample proportion of children who learn recorder.

(ii) Find an approximate 90% confidence interval for p.

(b) Mary now wants to construct an approximate 99% confidence interval for p such that the width of the confidence interval does not exceed 0.1. Using the result of (a)(i), estimate the least number of children that Mary should survey.

(6 marks)

2021 DSE Q10

The number of commercial emails that John receives each hour follows a Poisson distribution with a mean of 1.3 per hour, while the number of non-commercial emails that he receives each hour follows a Poisson distribution with a mean of 0.9 per hour.

(a) Find the probability that the number of non-commercial emails that John receives in a certain hour is fewer than 3. (3 marks)

(b) Find the probability that the number of commercial emails that John receives in 6 hours is 5. (2 marks)

(c) Find the probability that the number of emails that John receives in a certain hour is 2. (3 marks)

(d) Given that the number of emails that John receives in a certain hour is 2, find the probability that both emails are non-commercial emails. (3 marks)

(e) Given that the number of emails that John receives in a certain hour is fewer than 3, find the probability that John does not receive commercial email in that hour. (3 marks)

Provided by dse.life

11. Point and Interval Estimation

1. (2018 DSE-MATH-M1 Q2)

2. (2016 DSE-MATH-M1 Q4)

	Marks	Remarks
(a) The point estimate of p is $\dfrac{64}{100}=0.64$.	1A	
An approximate 95% confidence interval for p		
$=\left(\dfrac{64}{100}-1.96\sqrt{\dfrac{(0.64)(0.36)}{100}},\dfrac{64}{100}+1.96\sqrt{\dfrac{(0.64)(0.36)}{100}}\right)$	1M+1A	1A for 1.96
$=(0.54592,0.73408)$	1A	
$\approx(0.5459,0.7341)$		
(b) Let n be the number of packs in the sample.		
The width of a 90% confidence interval for p is $(2)(1.645)\left(\dfrac{0.05}{\sqrt{n}}\right)$.		
$(2)(1.645)\left(\dfrac{0.05}{\sqrt{n}}\right)<0.04$	1M+1A	
$\sqrt{n}>4.1125$		
$n>16.91265625$		
Thus, the least sample size is 17 .	1A	
	----------(7)	

(a)	Very good. More than 60% of the candidates were able to evaluate the approximate 95% confidence interval for the population proportion p . However, some candidates wrongly used $n=64$ instead of $n=100$ in evaluating the approximate confidence interval $\left(\dfrac{64}{100}-1.96\sqrt{\dfrac{(0.64)(0.36)}{n}},\dfrac{64}{100}+1.96\sqrt{\dfrac{(0.64)(0.36)}{n}}\right)$.
(b)	Good. Some candidates were unable to distinguish the concept between the confidence interval for the population mean and the approximate confidence interval for the population proportion.

3. (2014 DSE-MATH-M1 Q9)

	Marks
(a) An estimate of $p=\dfrac{80}{200}$	
$=0.4$	1A
An approximate 95% confidence interval for p	
$=\left(0.4-1.96\sqrt{\dfrac{0.4\times0.6}{200}},0.4+1.96\sqrt{\dfrac{0.4\times0.6}{200}}\right)$	1M
$\approx(0.3321,0.4679)$	1A
(b) $X\sim\mathrm{N}\left(0.85,\dfrac{0.85\,(1-0.85)}{n}\right)$	
$\mathrm{P}\left(X>\dfrac{100}{n}\right)<0.05$	1A
$\mathrm{P}\left(Z>\dfrac{\frac{100}{n}-0.85}{\sqrt{\frac{0.85(0.15)}{n}}}\right)<0.05$	1M
$\dfrac{100-0.85n}{n}\sqrt{\dfrac{n}{0.1275}}>1.645$	1M
$0.85n+1.645\sqrt{0.1275n}-100<0$	
$-11.19754391<\sqrt{n}<10.50650569$	
$0<n<110.3866618$	
Hence the maximum number of n is 110 .	1A
	(7)

(a) (b)	Very good. Poor. Some candidates considered $\mathrm{P}(X>100)$ rather than $\mathrm{P}(X>100/n)$. Some candidates used wrong means such as 85 and $0.85n$, or wrong standard deviations such as $\dfrac{0.85\,(1-0.85)}{n}$, for standardisation. Others got inequalities in $\sqrt{n}$ with incorrect direction of sign.

Provided by dse.life

4. (2013 DSE-MATH-M1 Q6)

An estimate for p is $\frac{75}{120}=0.625$. — 1A — OR $\frac{5}{8}$

An approximate 90% confidence interval for p

$\approx\left(0.625-1.645\sqrt{\frac{0.625(1-0.625)}{120}},\ 0.625+1.645\sqrt{\frac{0.625(1-0.625)}{120}}\right)$ — 1M+1M

$\approx(0.5523,\ 0.6977)$ — 1A

(4)

Good. Some candidates treated 75 as the sample size. Some wrote wrong expressions for the approximate standard deviation of the sample proportion.

5. (2013 DSE-MATH-M1 Q9)

(a) P (lifetime of a bulb <39000) $=0.9641$

$P\left(Z<\frac{39000-\mu}{5000}\right)=0.9641$ — 1M

$\frac{39000-\mu}{5000}\approx1.8$

$\mu\approx30000$ — 1A

(b) P ($30200<$ sample mean <30800) — Can use '$\le$' sign

$=P\left(\frac{30200-30000}{\frac{5000}{\sqrt{100}}}<Z<\frac{30800-30000}{\frac{5000}{\sqrt{100}}}\right)$ — 1M

$=P(0.4<Z<1.6)$

$\approx0.4452-0.1554$

$=0.2898$ — 1A

(c) P (sample mean >28500) ≥0.985

$P\left(Z>\frac{28500-30000}{\frac{5000}{\sqrt{n}}}\right)\ge0.985$ — 1M

$-0.3\sqrt{n}\le-2.17$ — 1A

$n\ge52.32111111$

Thus, the least value of n is 53. — 1A

(7)

(a)	Very good. Some candidates showed poor presentation such as '$P\left(Z<\frac{39000-\mu}{5000}\right)=0.9641$', '$P\left(\frac{39000-\mu}{5000}\right)=0.9641$', '$P\left(Z<\frac{39000-\mu}{5000}\right)=0.4641=0.9641$' or '$0.9641=0.18$'.
(b)	Good. Some candidates used '$0.4452+0.1554$' to find the probability required.
(c)	Fair. Some candidates used same symbols for both random variables before and after standardization. Many did not show enough ability to solve inequalities, such as writing a negative number greater than or equal to a positive number.

6. (2012 DSE-MATH-M1 Q6)

(a) Let X be the weight of a student. The sample mean $\overline{X}\sim N\left(67,\frac{15^2}{36}\right)$.

$P(\overline{X}>70)=P\left(Z>\frac{70-67}{\frac{15}{6}}\right)$ — 1M

$=P(Z>1.2)$

≈0.1151 — 1A

(b) The sample proportion is $\frac{9}{36}=0.25$. — 1A

An approximate 95% confidence interval for the proportion

$\approx\left(0.25-1.96\times\sqrt{\frac{0.25\times0.75}{36}},\ 0.25+1.96\times\sqrt{\frac{0.25\times0.75}{36}}\right)$ — 1M

$\approx(0.1085,\ 0.3915)$ — 1A

(5)

(a)	Good. Some candidates failed to perform the standardisation related to the distribution of a sample mean correctly.
(b)	Satisfactory. Many candidates found the sample proportion but failed to find the confidence interval required.

7. (PP DSE-MATH-M1 Q6)

(a) $Var(2\overline{X}+7)=4Var(\overline{X})$

$=4\left(\frac{8}{10}\right)$ — 1M — For $Var(\overline{X})=\frac{Var(X)}{n}$

$=3.2$ — 1A

(b) A 97% confidence interval for μ

$=\left(50-2.17\times\frac{\sqrt{8}}{\sqrt{10}},\ 50+2.17\times\frac{\sqrt{8}}{\sqrt{10}}\right)$ — 1M+1A — 1M for $50\pm d$; 1A for 2.17

$=(48.0591,\ 51.9409)$ — 1A

(5)

(a)	平平。 部分學生誤以爲 $Var(2\overline{X}+7)=2Var(\overline{X})$ 及 $Var(\overline{X})=Var(X)$。
(b)	平平。 部分學生誤以爲置信區間是 $\left(50-2.17\times\frac{8}{10},50+2.17\times\frac{8}{10}\right)$。

8. (SAMPLE DSE-MATH-M1 Q3)

(a) The estimate for the population proportion $=\frac{57}{150}=0.38$ — 1A

(b) The approximate 90% confidence interval for the population proportion

$=\left(0.38-1.645\sqrt{\frac{0.38(1-0.38)}{150}},\ 0.38+1.645\sqrt{\frac{0.38(1-0.38)}{150}}\right)$ — 1M+1M — 1M for using (a); 1M for $\overline{x}\pm z_c s$

$\approx(0.314805956,\ 0.445194043)$

$\approx(0.3148,\ 0.4452)$ — 1A

(4)

Provided by dse.life

9. (SAMPLE DSE-MATH-M1 Q5)

Solution	Marks	Remarks
Let X be the lifetime of the bulb, then by Central Limit Theorem, $\overline{X} \sim N\left(640, \frac{40^2}{25}\right)$ approximately.	1A	
The required probability is $P(\overline{X} > 630)$ $= P\left(Z > \frac{630-640}{\frac{40}{\sqrt{25}}}\right)$	1M+1A	1M for standardization
$= P(Z > -1.25)$ $= 0.5 + P(0 < Z < 1.25)$	1M	
$= 0.5 + 0.3944$ $= 0.8944$	1A	
	(5)	

Section B

10. (2017 DSE-MATH-M1 Q9)

Solution	Marks	Remarks
(a) (i) The sample mean $= \frac{(0.75)(11)+(1.25)(13)+(1.75)(8)+(2.25)(5)+(2.75)(3)}{40}$ $= 1.45$ hours	1A	
A 90% confidence interval for μ $= \left(1.45 - 1.645\left(\frac{0.4}{\sqrt{40}}\right), 1.45 + 1.645\left(\frac{0.4}{\sqrt{40}}\right)\right)$	1M+1A	1A for 1.645
$\approx (1.3460, 1.5540)$	1A	r.t. (1.3460 , 1.5540)
(ii) Let n be the sample size. $2(2.17)\left(\frac{0.4}{\sqrt{n}}\right) \le 0.3$	1M+1A	1A for 2.17
$n \ge 33.48551111$ Thus, the least sample size is 34 .	1A	
	(7)	
(b) (i) The required probability $= P\left(Z > \frac{2-1.48}{0.4}\right)$	1M	
$= P(Z > 1.3)$ $= 0.5 - 0.4032$ $= 0.0968$	1A	
(ii) The required probability $= \frac{C_1^9(1-0.0968)^8(0.0968)^2}{1-(1-0.0968)^{15} - C_1^{15}(1-0.0968)^{14}(0.0968)}$	1M+1M+1M	1M for using (b)(i) + 1M for numerator + 1M for denominator
[redacted] ≈ 0.0861	1A	r.t. 0.0861
	(6)	

(a) (i)	Very good. Most candidates were able to find the confidence interval correctly.
(ii)	Very good. A few candidates wrongly used the sample mean obtained in (a)(i) to find the width of the interval concerned.
(b) (i)	Very good. About 80% of the candidates were able to find the required probability.
(ii)	Good. Many candidates were able to find the required conditional probability.

Provided by dse.life

11. (2015 DSE-MATH-M1 Q9)

Solution	Marks	Remarks
(a) (i) The sample mean $= 68.64$ km/h	1A	
A 95% confidence interval for μ $= \left(68.64 - 1.96\left(\frac{16}{\sqrt{25}}\right), 68.64 + 1.96\left(\frac{16}{\sqrt{25}}\right)\right)$	1M+1A	1A for 1.96
$= (62.368, 74.912)$	1A	
(ii) Let n be the sample size. $2(2.24)\left(\frac{16}{\sqrt{n}}\right) < 9$	1M+1A	1A for 2.24
$n > 63.43237531$ Thus, the least sample size is 64.	1A	
	----------(7)	
(b) (i) The required probability $= P\left(Z > \frac{90-66}{16}\right)$	1M	
$= P(Z > 1.5)$ $= 0.5 - 0.4332$ $= 0.0668$	1A	
(ii) The required probability $= 1 - (1-0.0668)^{12} - C_1^{12}(1-0.0668)^{11}(0.0668) - C_2^{12}(1-0.0668)^{10}(0.0668)^2$	1M+1M	1M for using (b)(i) + 1M for binomial probability
≈ 0.041574551 ≈ 0.0416	1A	r.t. 0.0416
	----------(5)	

(a) (i)	Very good. Most candidates were able to use the correct formula to find the confidence interval while a few candidates treated 16 as the variance rather than the standard deviation of the given distribution.
(ii)	Very good. A few candidates wrongly used the sample mean in (a)(i) to find the width of the interval concerned.
(b) (i)	Very good. Most candidates were able to perform standardization and find the required probability.
(ii)	Very good. Most candidates were able to formulate the required probability form while a few candidates used wrong probabilities in substitution.

12. (2014 DSE-MATH-M1 Q12)

Solution	Marks	Remarks
(a) $P(\mu - 3.5 \le X \le \mu + 3.5) = 0.5160$	1A	
$P\left(0 \le Z \le \frac{3.5}{\sigma}\right) = 0.2580$ $\frac{3.5}{\sigma} = 0.7$ $\sigma = 5$	1A	
$P(X > 25) = 0.2743$	1A	
$P\left(0 < Z < \frac{25-\mu}{\sigma}\right) = 0.2257$ $\frac{25-\mu}{5} = 0.6$ $\mu = 22$	1A	
	(4)	
(b) $P(X > k) \le \frac{5}{200}$	1A	
$P\left(0 < Z < \frac{k-22}{5}\right) \ge 0.475$ $\frac{k-22}{5} \ge 1.96$	1M	
$k \ge 31.8$ Hence the minimum integral value of k is 32.	1A	
	(3)	
(c) (i) Sample mean $= \frac{22+15+\cdots+24}{12}$ $= 21$	1A	
A 90% confidence interval $\approx \left(21 - 1.645 \times \frac{4.7}{\sqrt{12}}, 21 + 1.645 \times \frac{4.7}{\sqrt{12}}\right)$	1M	
$\approx (18.7681, 23.2319)$	1A	
(ii) Let $\overline{Y}$ be the mean delivery time of the n orders. $P(\theta - 3 \le \overline{Y} \le \theta + 3) > 0.99$		Alternative Solution $\left(\overline{Y} - 2.575 \times \frac{4.7}{\sqrt{n}}, \overline{Y} + 2.575 \times \frac{4.7}{\sqrt{n}}\right)$
$P\left(\frac{-3}{\frac{4.7}{\sqrt{n}}} \le Z \le \frac{3}{\frac{4.7}{\sqrt{n}}}\right) > 0.99$	1M	$\subsetneqq (\overline{Y} - 3, \overline{Y} - 3)$
$\frac{3}{\frac{4.7}{\sqrt{n}}} > 2.575$	1M	$\therefore 2.575 \times \frac{4.7}{\sqrt{n}} < 3$
$n > 16.27450069$ Hence the minimum value of n is 17.	1A	
	(6)	

(a)	Satisfactory. Some candidates found the value of $P(\mu - 1.75 \le X \le \mu + 1.75)$ instead of $P(\mu - 3.5 \le X \le \mu + 3.5)$.
(b)	Fair. Some candidates got inequalities in k with incorrect direction of sign.
(c) (i)	Good. Few candidates used the sample standard deviation in the calculation.
(ii)	Poor. Some candidates wrongly treated 21 as the mean of $\overline{Y}$. Some candidates thought that the length of the confidence interval was 3.

Provided by dse.life

13. (2013 DSE-MATH-M1 Q12)

(a) (i) $2\times 1.96\times\frac{\sigma}{\sqrt{49}}=5.044-4.596$ — 1M

$\sigma=0.8$ — 1A

(ii) The mean of the sample $=\frac{4.596+5.044}{2}$

$=4.82$ — 1A

(3)

(b) The combined sample mean $=\frac{4.82\times 49+3.6+3.8+\cdots+6.4}{49+15}$ — 1M

$=4.83875$ — 1A — OR 4.8388

A 99% confidence interval for μ

$\approx\left(4.83875-2.575\times\frac{0.8}{\sqrt{64}},\ 4.83875+2.575\times\frac{0.8}{\sqrt{64}}\right)$ — 1M

$=(4.58125,\ 5.09625)$ — 1A — OR (4.5813, 5.0963)

(4)

(c) Let X be the cholesterol level of a randomly selected adult.

(i) P(low) $=P(X\le 5.2)$

$=P\left(Z\le\frac{5.2-4.8}{0.8}\right)$

$=P(Z\le 0.5)$

≈ 0.6915 — 1A

(ii) P(high) $=P(X\ge 6.2)$

$=P\left(Z\ge\frac{6.2-4.8}{0.8}\right)$

$=P(Z\ge 1.75)$

≈ 0.0401 — 1A

P(medium) $\approx 1-0.6915-0.0401$

$=0.2684$ — 1A

P(more than 17 adults with low level and at least 1 adult with medium level)

$\approx C^{20}_{18}(0.6915)^{18}[C^2_1(0.2684)(0.0401)+(0.2684)^2]+C^{20}_{19}(0.6915)^{19}(0.2684)$ — 1M

≈ 0.0281 — 1A

(5)

(a)	Good. Some wrote $\frac{\sigma}{49}$ or $\frac{\sigma^2}{\sqrt{49}}$ instead of $\frac{\sigma}{\sqrt{49}}$ as the standard deviation of the sample mean.
(b)	Poor. Many candidates tried to use the standard deviation of the combined sample instead of that of the population. Some thought that the mean of the combined sample would be $\frac{4.82+4.9}{2}$.
(c) (i)	Fair. Some candidates assigned wrong values to μ and σ.
(ii)	Poor. Some candidates missed out the factor C^2_1 in calculating the probability required.

14. (2012 DSE-MATH-M1 Q12)

(a) (i) The sample mean $=\frac{56+\cdots+50}{16}$

$=51.5625$ — 1A

A 90% confidence interval

$\approx\left(51.5625-1.645\times\frac{9}{\sqrt{16}},\ 51.5625+1.645\times\frac{9}{\sqrt{16}}\right)$ — 1M+1A

$=(47.86125,\ 55.26375)$ — 1A — OR (47.8613, 55.2638)

(ii) Let n be the sample size.

$\therefore\ 2\left(1.645\cdot\frac{9}{\sqrt{n}}\right)<6$ — 1M

$n>24.354225$ — 1A

Hence, the least sample size is 25. — 1A

(7)

(b) (i) P(a tourist waits for more than 65 minutes)

$=P\left(Z>\frac{65-51.5}{9}\right)$ — 1M

$=P(Z>1.5)$

≈ 0.0668 — 1A

P(less than 2 coupons are sent to the first 10 tourists interviewed)

$\approx(1-0.0668)^{10}+C^{10}_1(1-0.0668)^9(0.0668)$ — 1M

≈ 0.8594 — 1A

(ii) P(the 5th coupon is sent to the 20th tourist interviewed)

$\approx C^{19}_4(1-0.0668)^{15}(0.0668)^4\cdot 0.0668$ — 1M

≈ 0.0018 — 1A

(6)

(a) (i)	Good. However, some candidates used the standard deviation of the sample instead of the population, used values other than 1.645, or interchanged the upper and lower confidence limits.
(ii)	Fair. Besides mistakes similar to (i), many candidates did not write the width of the confidence interval correctly or failed to solve inequalities.
(b) (i)	Good. Most candidates were able to express the probability of the mentioned event, but some failed in the standardisation of normal distributions.
(ii)	Satisfactory. Binomial coefficients were omitted or written wrongly by some candidates.

Provided by dse.life

15. (PP DSE-MATH-M1 Q12)

Solution	Marks	Remarks
(a) The estimate of the mean $= \dfrac{0\times 6+\cdots+7\times 4}{100}$ $= 3.21$	1A	
	(1)	
(b) (i) The sample proportion of school days with less than 4 visits $= \dfrac{57}{100}$	1A	
(ii) An approximate 95% confidence interval for the proportion $= \left(0.57-1.96\sqrt{\dfrac{0.57\times 0.43}{100}},\ 0.57+1.96\sqrt{\dfrac{0.57\times 0.43}{100}}\right)$	1M	
$=(0.4730, 0.6670)$	1A	
	(3)	
(c) (i) By (a), $\lambda = 3.21$. P(*crowded* on a day) $= 1-e^{-3.21}\left(1+3.21+\dfrac{3.21^2}{2!}+\dfrac{3.21^3}{3!}\right)$	1M	For Poisson probability
≈ 0.399705729 ≈ 0.3997	1A	
(ii) P(*crowded* on alternate days | *crowded* on at least 2 days) $= \dfrac{(0.399705729)^3(1-0.399705729)^2+(1-0.399705729)^3(0.399705729)^2}{1-(1-0.399705729)^5-5(1-0.399705729)^4(0.399705729)}$	1M+1M+1M	1M for numerator 1M for denominator 1M for binomial probability
≈ 0.0869	1A	
	(6)	

(a)	良好。 少數學生誤以爲估算值是 $\dfrac{321}{99}$ 。
(b) (i)	良好。
(ii)	平平。 部分學生未能運用總體比例的置信區間公式。
(c) (i)	平平。 少部分學生誤以爲所求概率是 $1-P(0)-P(1)-P(2)$ 。
(ii)	甚差。 學生在研究所有的可能性時出現困難。

16. (SAMPLE DSE-MATH-M1 Q14)

Solution	Marks	Remarks
(a) (i) An unbiased estimate for μ is 18.6.	1A	
(ii) The 95% confidence interval for μ $= \left(18.6-1.96\dfrac{4.5}{\sqrt{9}},\ 18.6+1.96\dfrac{4.5}{\sqrt{9}}\right)$	1M	1M for $\bar{x}\pm z_c s$
$=(15.66, 21.54)$	1A	
	(3)	
(b) Let $\bar{X}$ be the sample mean of the BMI values of 25 children aged 12, then $\bar{X}\sim N\left(18.7, \dfrac{4.5^2}{25}\right)$.	1A	
The required probability $= P(\bar{X}<17.8)$ $= P\left(Z<\dfrac{17.8-18.7}{4.5/\sqrt{25}}\right)$	1M+1M	1M for standardization 1M for $\dfrac{\sigma}{\sqrt{n}}$
$= P(Z<-1)$ $= 0.5-0.3413$ $= 0.1587$	1A	
	(4)	
(c) (i) Let X be the BMI values of the children aged 12, then $X\sim N(18.7, 4.5^2)$. The required probability $= P(X>25)$ $= P\left(Z>\dfrac{25-18.7}{4.5}\right)$	1M	For standardization
$= P(Z>1.4)$ $= 0.5-0.4192$ $= 0.0808$	1A	
(ii) The required probability $= (1-0.0808)^8 + {}_8C_1(1-0.0808)^7(0.0808)$	1M+1M	1M for 2 cases 1M for binomial prob
≈ 0.868062354 ≈ 0.8681	1A	
(iii) The required probability $= \dfrac{{}_9C_1(1-0.0808)^8(0.0808)\times(0.0808)}{(1-0.0808)^8+{}_8C_1(1-0.0808)^7(0.0808)}$	1M+1M	1M for numerator 1M for denominator using (c)(ii)
≈ 0.034498041 ≈ 0.0345	1A	
<u>Alternative solution</u> The required probability $= \dfrac{(1-0.0808)^8\times(0.0808)^2+{}_8C_1(1-0.0808)^7(0.0808)\times(1-0.0808)(0.0808)}{(1-0.0808)^8+{}_8C_1(1-0.0808)^7(0.0808)}$	1M+1M	1M for numerator 1M for denominator using (c)(ii)
≈ 0.034498041 ≈ 0.0345	1A	
	(8)	

Provided by dse.life