Solutionbank S3

Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 1

Question:

The time, in minutes, it takes Robert to complete the puzzle in his morning newspaper each day is normally distributed with mean 18 and standard deviation 3. After taking a holiday, Robert records the times taken to complete a random sample of 15 puzzles and he finds that the mean time is 16.5 minutes. You may assume that the holiday has not changed the standard deviation of times taken to complete the puzzle.
Stating your hypotheses clearly test, at the 5% level of significance, whether or not there has been a reduction in the mean time Robert takes to complete the puzzle. ***E***

Solution:

$H_0: \mu = 18, H_1: \mu < 18$

Remember to identify which is H_0 and which is H_1. This is a one-tail test since we are only interested in whether the time taken to solve the puzzle has reduced. You must use the correct parameter (μ).

$$z = \frac{(16.5 - 18)}{\left(\frac{3}{\sqrt{15}}\right)} = -1.9364\ldots$$

Using $z = \frac{x - \mu}{\frac{\sigma}{\sqrt{n}}}$.

5% one tail c.v. is $z = -1.6449$

Use the percentage point table and quote the figure in full.

significant *or* reject H_0 *or* in critical region.

$-1.9364 < -1.6449$

There is evidence that the (mean) time to complete the puzzles has reduced

or Robert is getting faster (at doing the puzzles)

State your conclusion in the context of the question.

Solutionbank S3
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 2

Question:

In a trial of diet A a random sample of 80 participants were asked to record their weight loss, x kg, after their first week of using the diet. The results are summarised by

$$\sum x = 361.6 \quad \text{and} \quad \sum x^2 = 1753.95.$$

a Find unbiased estimates for the mean and variance of weight lost after the first week of using diet A.

The designers of diet A believe it can achieve a greater mean weight loss after the first week than a standard diet B. A random sample of 60 people used diet B. After the first week they had achieved a mean weight loss of 4.06 kg, with an unbiased estimate of variance of weight loss of $2.50\,\text{kg}^2$.

b Test, at the 5% level of significance, whether or not the mean weight loss after the first week using diet A is greater than that using diet B. State your hypotheses clearly.

c Explain the significance of the Central Limit Theorem to the test in part **b**.

d State an assumption you have made in carrying out the test in part **b**. ***E***

Solution:

a $\bar{x} = \frac{361.6}{80} = 4.52$

$\hat{\sigma}^2 = s^2 = \frac{1753.95 - 80 \times \bar{x}^2}{79} = 1.5128$

or $\hat{\sigma}^2 = s^2 = \frac{80}{79} \times \left(\frac{1753.95}{80} - \bar{x}^2\right) = 1.5128$

Using $\frac{\sum x^2 - n\bar{x}^2}{n-1}$ or $\frac{n}{n-1}\left(\frac{\sum x^2}{n} - \bar{x}^2\right)$.

b $H_0: \mu_A = \mu_B \quad H_1: \mu_A > \mu_B$

This is a difference of means test. When stating hypotheses you must make it clear which mean is greater when it is one-tailed test.

$z = \frac{4.52 - 4.06}{\sqrt{\frac{1.5128}{80} + \frac{2.50}{60}}} = \left(\frac{0.46}{\sqrt{0.060576}}\right)$

$= 1.8689$ or -1.8689 if $B - A$ was used.

Using $z = \frac{(\bar{A} - \bar{B}) - (\mu_A - \mu_B)}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}}$.

One tail c.v. is $z = 1.6449$

$1.87 > 1.6449$ so reject H_0.

Use the percentage point table and quote the figure in full.

There is evidence that diet A is better than diet B *or* evidence that (mean) weight lost in first week using diet A is more than with B.

State your conclusion in the context of the question.

c CLT enables you to assume that $\bar{A}$ and $\bar{B}$ are normally distributed since both samples are large.

d Assumed $\sigma_A^2 = s_A^2$ and $\sigma_B^2 = s_B^2$

Variance must be known to use the test. Remember σ^2 is the population variance and s^2 is an unbiased estimator of the population variance.

Solutionbank S3

Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 3

Question:

A random sample of the daily sales (in £s) of a small company is taken and, using tables of the normal distribution, a 99% confidence interval for the mean daily sales is found to be

(123.5,154.7).

Find a 95% confidence interval for the mean daily sales of the company. ***E***

Solution:

$$123.5 = \bar{x} - 2.5758 \times \frac{\sigma}{\sqrt{n}} \quad ①$$

$$154.7 = \bar{x} + 2.5758 \times \frac{\sigma}{\sqrt{n}} \quad ②$$

99% confidence interval so each tail is 0.05. Use the percentage point table and quote the figure in full. C.I. $\bar{x} \pm 2.5758 \times \frac{\sigma}{\sqrt{n}}$

$$\bar{x} = \frac{1}{2}(123.5 + 154.7) = 139.1$$

Add equations ①and ② to find $\bar{x}$ or calculate the mean of the given limits.

$$2.5758 \times \frac{\sigma}{\sqrt{n}} = 154.7 - 139.1$$

$$= 15.6$$

$$\frac{\sigma}{\sqrt{n}} = \frac{15.6}{2.5758}$$

Substitute $\bar{x}$ into equation ① or ② to find $\frac{\sigma}{\sqrt{n}}$.

$$\text{So 95\% C.I.} = 139.1 \pm 1.9600 \times \frac{15.6}{2.5758}$$

$$= (127.22\ldots, 150.97\ldots)$$

95% confidence interval so each tail is 0.025. Substitute in $\bar{x}$ and $\frac{\sigma}{\sqrt{n}}$.

$$= (127, 151)$$

Answers should be given to at least 3 significant figures.

Solutionbank S3
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 4

Question:

A set of scaffolding poles come in two sizes, long and short. The length L of a long pole has the normal distribution $N(19.7, 0.5^2)$. The length S of a short pole has the normal distribution $N(4.9, 0.2^2)$. The random variables L and S are independent. A long pole and a short pole are selected at random.

a Find the probability that the length of the long pole is more than 4 times the length of the short pole.

Four short poles are selected at random and placed end to end in a row. The random variable T represents the length of the row.

b Find the distribution of T.

c Find $P(|L - T| < 0.1)$. ***E***

Solution:

a Let $X = L - 4S$ then
$E(X) = 19.7 - 4 \times 4.9 = 0.1$

Use $E(aX - bY) = aE(X) - bE(X)$.

$$\text{Var}(X) = \text{Var}(L) + 4^2\text{Var}(S) = 0.5^2 + 16 \times 0.2^2 = 0.89$$

Use $\text{Var}(aX - bY) = a^2\text{Var}(X) + b^2\text{Var}(Y)$.

$$P(X > 0) = P\left(Z > \frac{0 - 0.1}{\sqrt{0.89}}\right) = P(Z > -0.10599\ldots) = (0.542 \text{ to } 0.544)$$

Use $z = \frac{x - \mu}{\sigma}$.

Look z up in the tables or use a graphical calculator.

This is the accepted range of values for the answer.

b $T = S_1 + S_2 + S_3 + S_4$

$$E(T) = 4 \times 4.9 = 19.6$$
$$\text{Var}(T) = 4 \times 0.2^2 = 0.16 \text{ or } 0.4^2$$

Since we are adding we use $E(X_1 + X_2) = E(X_1) + E(X_2)$ and $\text{Var}(X_1 + X_2) = \text{Var}(X_1) + \text{Var}(X_2)$.

c Let $Y = L - T$

$$E(Y) = E(L) - E(T) = 0.1$$

Using $E(X - Y) = E(X) - E(Y)$.

$$\text{Var}(Y) = \text{Var}(L) + \text{Var}(T) = 0.5^2 + 0.4^2 = 0.41$$

Using $\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y)$.

$$P(-0.1 < Y < 0.1) = P\left(Z < \frac{0.1 - 0.1}{\sqrt{0.41}}\right) - P\left(Z < \frac{0.1 - 0.1}{\sqrt{0.41}}\right)$$
$$= P(Z < 0) - P(Z < -0.31)$$
$$= 0.5 - (1 - 0.6217)$$
$$= 0.1217 \text{ (tables)}$$
$$\text{or } 0.1226 \ldots \text{ (calc)}$$

You do not interpolate so round your z value to 2 decimal places and use the tables, or use a graphical calculator.

Solutionbank S3

Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 5

Question:

Describe one advantage and one disadvantage of

a quota sampling,

b simple random sampling. ***E***

Solution:

a *Advantages*: Any one of

- does not require the existence of:
  a sampling frame
  a population list
- field work can be done quickly as representative sample can be achieved with a small sample size
- costs kept to a minimum (cheaply)
- administration relatively easy
- non-response not an issue

This is a book work. You need to learn the advantages and disadvantages of all the sampling methods.

Disadvantages: Any one of

- not possible to estimate sampling errors
- interviewer choice and may not be able to judge easily/may lead to bias
- non-response not recorded
- non-random process

b *Advantages*: any one of

- random process so possible to estimate sampling errors
- free from bias

Disadvantages: any one of

- not suitable when sample size is large
- sampling frame required which may not exist or may be difficult to construct for a large population

Solutionbank S3

Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 6

Question:

A report on the health and nutrition of a population stated that the mean height of three-year-old children is 90 cm and the standard deviation is 5 cm. A sample of 100 three-year-old children was chosen from the population.

a Write down the approximate distribution of the sample mean height. Give a reason for your answer.

b Hence find the probability that the sample mean height is at least 91 cm. ***E***

Solution:

a $\bar{X} \sim \text{N}\left(90, \frac{5^2}{100}\right)$ i.e. N (90, 0.25)

Application of Central Limit Theorem (as sample large)

b $\text{P}(\bar{X} \geq 91) = 1 - \text{P}\left(Z < \frac{91-90}{0.25}\right)$

$= 1 - \text{P}(Z < 2)$

$= 1 - 0.9772$

$= 0.0228$

Using $z = \frac{x - \mu}{\frac{\sigma}{\sqrt{n}}}$.

You do not interpolate so round your z value to 2 decimal places and use the tables or use a graphical calculator.

Solutionbank S3
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 7

Question:

A machine produces metal containers. The weights of the containers are normally distributed. A random sample of 10 containers from the production line was weighed, to the nearest 0.1 kg, and gave the following results

49.7, 50.3, 51.0, 49.5, 49.9
50.1, 50.2, 50.0, 49.6, 49.7.

a Find unbiased estimates of the mean and variance of the weights of the population of metal containers.

The machine is set to produce metal containers whose weights have a population standard deviation of 0.5 kg.

b Estimate the limits between which 95% of the weights of metal containers lie.

c Determine the 99% confidence interval for the mean weight of metal containers. ***E***

Solution:

a $\bar{X} = \frac{500}{10} = 50$

$s^2 = \frac{25\,001.74 - 10 \times 50^2}{9}$

$= 0.193$

Using $\frac{\sum x^2 - n\bar{x}^2}{n-1}$ or $\frac{n}{n-1}\left(\frac{\sum x^2}{n} - \bar{x}^2\right)$.

b Limits are $50 \pm 1.96 \times 0.5$
$= (49.02, 50.98)$

c Confidence interval is

$(50 - 2.5758 \times \frac{0.5}{\sqrt{10}}, 50 + 2.5758 \times \frac{0.5}{\sqrt{10}})$
$= (49.593, 50.407)$
$= (49.6, 50.4)$

99% confidence interval so each tail is 0.005. Use the percentage point table and quote the figure in full.
C.I.: $\bar{x} \pm 2.5758 \times \frac{\sigma}{\sqrt{n}}$

Final answers should be given to at least 3 significant figures.

Solutionbank S3

Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 8

Question:

A school has 15 classes and a sixth form. In each class there are 30 students. In the sixth form there are 150 students. There are equal numbers of boys and girls in each class. There are equal numbers of boys and girls in the sixth form. The headteacher wishes to obtain the opinions of the students about school uniforms.
Explain how the headteacher would take a stratified sample of size 40. ***E***

Solution:

Total in school $= (15 \times 30) + 150 = 600$

Need a random sample of $\frac{30}{600} \times 40$

$= 2$ from each of the 15 classes

and a random sample of $\frac{150}{600} \times 40$

$= 10$ from sixth form;

To describe a stratified sample your need to
1) work out how many people to take from each strata,
2) explain how to collect the sample from each strata.

Label the boys in each class from 1–15 and the girls from 1–15.
Use random numbers to select 1 girl and 1 boy from each class.
Label the boys in the sixth form from 1–75 and the girls from 1–75.
Use random numbers to select 5 different boys and 5 different girls from the sixth form.

Solutionbank S3
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 9

Question:

A workshop makes two types of electrical resistor.
The resistance, X ohms, of resistors of Type A is such that $X \sim N(20, 4)$.
The resistance, Y ohms, of resistors of Type B is such that $Y \sim N(10, 0.84)$.
When a resistor of each type is connected into a circuit, the resistance R ohms of the circuit is given by $R = X + Y$ where X and Y are independent.
Find

a $E(R)$,

b $\mathrm{Var}(R)$,

c $P(28.9 < R < 32.64)$. ***E***

Solution:

a $E(R) = 20 + 10 = 30$

Using $E(X+Y) = E(X) + E(Y)$.

b $\mathrm{Var}(R) = 4 + 0.84$
$= 4.84$

Using $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

c $R \sim N(30, 4.84)$

$P(28.9 < R < 32.64) = P(R < 32.64) - P(R < 28.9)$

$= P\left(Z < \frac{32.64 - 30}{2.2}\right) - P\left(Z < \frac{28.9 - 30}{2.2}\right)$

Using $z = \frac{x - \mu}{\sigma}$.

$= P(Z < 1.2) - P(Z < -0.5)$

$= 0.8849 - (1 - 0.6915)$

$= 0.8849 - 0.3085$

$= 0.5764$

Look z up in the tables or use a graphical calculator.

Solutionbank S3
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 10

Question:

The drying times of paint can be assumed to be normally distributed. A paint manufacturer paints 10 test areas with a new paint. The following drying times, to the nearest minute, were recorded.

82, 98, 140, 110, 90, 125, 150, 130, 70, 110.

a Calculate unbiased estimates for the mean and the variance of the population of drying times of this paint.

Given that the population standard deviation is 25,

b find a 95% confidence interval for the mean drying time of this paint.

Fifteen similar sets of tests are done and the 95% confidence interval is determined for each set.

c Estimate the expected number of these 15 intervals that will enclose the true value of the population mean μ. ***E***

Solution:

a $\hat{\mu} = \bar{x} = \dfrac{82+98+140+110+90+125+150+130+70+110}{10}$

$= 110.5$

$s^2 = \dfrac{128\,153 - 10 \times 110.5^2}{9}$

$= 672.28$

Using $\dfrac{\sum x^2 - n\bar{x}^2}{n-1}$ or $\dfrac{n}{n-1}\left(\dfrac{\sum x^2}{n} - \bar{x}^2\right)$.

b 95% confidence limits are

$110.5 \pm 1.96 \times \dfrac{25}{\sqrt{10}}$

$= (95.005, 125.995)$

$= (95.0, 126)$

95% confidence interval so each tail is 0.025. Use the percentage point table and quote the figure in full.

C.I.: $\bar{x} \pm 1.9600 \times \dfrac{\sigma}{\sqrt{n}}$

Answers should be given to at least 3 significant figures.

c Number of intervals $= \dfrac{95}{100} \times 15 = 14.25$

Solutionbank S3

Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 11

Question:

Some biologists were studying a large group of wading birds. A random sample of 36 were measured and the wing length, x mm, of each wading bird was recorded. The results are summarised as follows.

$$\sum x = 6046, \sum x^2 = 1\ 06\ 338.$$

a Calculate unbiased estimates of the mean and the variance of the wing lengths of these birds.

Given that the standard deviation of the wing lengths of this particular type of bird is actually 5.1 mm,

b find a 99% confidence interval for the mean wing length of the birds from this group. ***E***

Solution:

a $\bar{x} = \left(\frac{6046}{36} = \right) 167.94\ldots$

$$s^2 = \frac{1\ 016\ 338 - 36 \times \bar{x}^2}{35}$$

$$= 27.0$$

Using $\frac{\sum x^2 - n\bar{x}^2}{n-1}$ or $\frac{n}{n-1}\left(\frac{\sum x^2}{n} - \bar{x}^2\right)$.

b 99% confidence interval is: $\bar{x} \pm 2.5758 \times \frac{5.1}{\sqrt{36}}$

99% confidence interval so each tail is 0.005. Use the percentage point table and quote the figure in full.

C.I.: $\bar{x} \pm 2.5758 \times \frac{\sigma}{\sqrt{n}}$

$$= 167.94 \pm 2.5758 \times \frac{5.1}{\sqrt{36}}$$

$$= (165.75, 170.13)$$

$$= (166, 170)$$

Answers should be given to at least 3 significant figures.

Solutionbank S3
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 12

Question:

The weights of adult men are normally distributed with a mean of 84 kg and a standard deviation of 11 kg.

a Find the probability that the total weight of 4 randomly chosen adult men is less than 350 kg.

The weights of adult women are normally distributed with a mean of 62 kg and a standard deviation of 10 kg.

b Find the probability that the weight of a randomly chosen adult man is less than one and a half times the weight of a randomly chosen adult woman. ***E***

Solution:

a $X = M_1 + M_2 + M_3 + M_4$

$E(X) = 4 \times 84$

$= 336$

$\text{Var}(X) = 4 \times 11^2$

$= 484 \text{ or } 22^2$

$X \sim N(336, 22^2)$

$P(X < 350) = P\left(Z < \frac{350 - 336}{22}\right)$

$= P(Z < 0.64)$

$= 0.738 \text{ or } 0.739$

Since we are adding we use $E(X_1 + X_2) = E(X_1) + E(X_2)$ and $\text{Var}(X_1 + X_2) = \text{Var}(X_1) + \text{Var}(X_2)$.

Using $z = \frac{x - \mu}{\sigma}$.

You do not interpolate so round your z value to 2 decimal places and use the tables, or use a graphical calculator.

b $M \sim N(84, 121)$ and $W \sim N(62, 100)$

Let $Y = M - 1.5W$

$E(Y) = 84 - 1.5 \times 62 = -9$

$\text{Var}(Y) = \text{Var}(M) + 1.5^2\text{Var}(W)$

$= 11^2 + 1.5^2 \times 10^2$

$= 346$

$P(Y < 0) = P\left(Z < \frac{0 - (-9)}{\sqrt{346}}\right)$

$= P(Z < 0.48)$

$= 0.684 \sim 0.686$

Using $E(X_1 - bX_2) = E(X_1) - bE(X_2)$ and $\text{Var}(X_1 - bX_2) = \text{Var}(X_1) + b^2\text{Var}(X_2)$.

Using $z = \frac{x - \mu}{\sigma}$.

You do not interpolate so round your z value to 2 decimal places and use the tables or use a graphical calculator.

Solutionbank S3
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 13

Question:

A researcher is hired by a cleaning company to survey the opinions of employees on a proposed pension scheme. The company employs 55 managers and 495 cleaners.
To collect data the researcher decides to give a questionnaire to the first 50 cleaners to leave at the end of the day.

a Give 2 reasons why this method is likely to produce biased results.

b Explain briefly how the researcher could select a sample of 50 employees using

- **i** a systematic sample,
- **ii** a stratified sample.

Using the random number tables in the formulae book, and starting with the top left hand corner (8) and working across, 50 random numbers between 1 and 550 inclusive were selected. The first two suitable numbers are 384 and 100.

c Find the next two suitable numbers. ***E***

Solution:

a Only cleaners – no managers i.e. not all types.
or not a random sample; 1st 50 may be in same shift/group/share same views.

b **i** Label employees (1–550) or obtain an ordered list.
Select first using random numbers (from 1–11).
Then select every 11th person from the list
e.g. if person 8 is selected then the sample is 8, 19, 30, 41, ... 547

ii Label managers (1–55) and cleaners (1–495)
Use random numbers to select...

iii ...5 managers and 45 cleaners

c 390, 372 (They must be in this order.)

Solutionbank S3
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 14

Question:

A sociologist is studying how much junk food teenagers eat. A random sample of 100 female teenagers and an independent random sample of 200 male teenagers were asked to estimate what their weekly expenditure on junk food was. The results are summarised below.

	n	mean	s.d.
Female teenagers	100	£5.48	£3.62
Male teenagers	200	£6.86	£4.51

a Using a 5% significance level, test whether or not there is a difference in the mean amounts spent on junk food by male teenagers and female teenagers. State your hypotheses clearly.

b Explain briefly the importance of the Central Limit Theorem in this problem. ***E***

Solution:

a $H_0 : \mu_F = \mu_M \quad H_1 : \mu_F \neq \mu_M$

This is a difference of means test.

$$z = \frac{6.86 - 5.48}{\sqrt{\frac{4.51^2}{200} + \frac{3.62^2}{100}}}$$

$= 2.860\ldots$

Using $z = \dfrac{(\bar{A} - \bar{B}) - (\mu_A - \mu_B)}{\sqrt{\dfrac{\sigma_A^2}{n_A} - \dfrac{\sigma_B^2}{n_B}}}$.

2-tail 5% critical value $(\pm)$ 1.96

$2.860 > +1.900$

Use the percentage point table and quote the figure in full.

Significant result or reject the null hypothesis.
There is evidence of a difference in the (mean) amount spent on junk food by male and female teenagers.

State your conclusion in the context of the question.

b CLT enables us to assume $\bar{F}$ and $\bar{M}$ are normally distributed.

Solutionbank S3
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 15

Question:

a State two reasons why stratified sampling might be chosen as a method of sampling when carrying out a statistical survey.
b State one advantage and one disadvantage of quota sampling. ***E***

Solution:

a Population divides into mutually exclusive/distinct groups/strata.
Its results will best reflect those of the population since the sample structure reflects that of the population.

b *Advantages*: Any one of
- enables fieldwork to be done quickly
- costs kept to a minimum
- administration is relatively easy

Disadvantages: Any one of
- non-random so not possible to estimate sampling errors
- subject to possible interviewer bias
- non-response not recorded

Solutionbank S3

Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 16

Question:

A sample of size 5 is taken from a population that is normally distributed with mean 10 and standard deviation 3. Find the probability that the sample mean lies between 7 and 10. ***E***

Solution:

$X \sim N(10, 3^2) \therefore \bar{X} \sim N(10, \frac{9}{5})$

$$P(7 \leq \bar{X} \leq 10) = P\left(\frac{7-10}{\sqrt{\frac{9}{5}}} < Z < 0\right)$$

$= P(-2.236 < Z < 0)$

$= \Phi(0) - \{1 - \Phi(2.24)\}$

$= 0.4875$

Using $z = \dfrac{x-\mu}{\frac{\sigma}{\sqrt{n}}}$

You do not interpolate so round your z value to 2 decimal places and use the tables, or use a graphical calculator.

Solutionbank S3

Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 17

Question:

A computer company repairs large numbers of PCs and wants to estimate the mean time to repair a particular fault. Five repairs are chosen at random from the company's records and the times taken, in seconds, are

205 310 405 195 320.

a Calculate unbiased estimates of the mean and the variance of the population of repair times from which this sample has been taken.

It is known from previous results that the standard deviation of the repair time for this fault is 100 seconds. The company manager wants to ensure that there is a probability of at least 0.95 that the estimate of the population mean lies within 20 seconds of its true value.

b Find the minimum sample size required. ***E***

Solution:

a Let X represent repair time

$\therefore \sum x = 1435 \therefore \bar{x} = \frac{1435}{5} = 287$

$\sum x^2 = 442\,575$

$\therefore s^2 = \frac{442\,575 - 5 \times 287^2}{4}$

$= 7682.5$

Using $\frac{\sum x^2 - n\bar{x}^2}{n-1}$ or $\frac{n}{n-1}\left(\frac{\sum x^2}{n} - \bar{x}^2\right)$.

b $P(|\mu - \hat{\mu}|) < 20 = 0.95$

The repair time is between 80 and 120. 95% confidence interval so each tail is 0.025. Use the percentage point table and quote the figure in full.
C.I.: $\bar{x} \pm 1.96 \times \frac{\sigma}{\sqrt{n}}$.

$\therefore 1.96 \times \frac{\sigma}{\sqrt{n}} = 20$

$\therefore n = \frac{1.96^2\sigma^2}{20^2} = \frac{1.96^2\sigma^2}{400} = \frac{1.96^2 \times 100^2}{400} = 96.04$

$\therefore$ Sample size $(\geq)$ 97 required

Solutionbank S3
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 18

Question:

The random variable D is defined as

$$D = A - 3B + 4C$$

where $A \sim \text{N}(5, 2^2)$, $B \sim \text{N}(7, 3^2)$ and $C \sim \text{N}(9, 4^2)$, and A, B and C are independent.

a Find $\text{P}(D < 44)$.

The random variables B_1, B_2 and B_3 are independent and each has the same distribution as B. The random variable X is defined as

$$X = A - \sum_{i=1}^{3} B_i + 4C.$$

b Find $\text{P}(X > 0)$. ***E***

Solution:

a
$$\text{E}(D) = \text{E}(A) - 3\text{E}(B) + 4\text{E}(C)$$
$$= 5 - 3 \times 7 + 4 \times 9$$
$$= 20$$

Using $\text{E}(aX \pm bY) = a\text{E}(X) \pm b\text{E}(Y)$

$$\text{Var}(D) = \text{Var}(A) + 9\text{Var}(B) + 16\text{Var}(C)$$
$$= 2^2 + 9 \times 3^2 + 16 \times 4^2$$
$$= 341$$

Using $\text{Var}(aX \pm bY) = a^2\text{Var}(X) \pm b^2\text{Var}(Y)$.

$$\text{P}(D < 44) = \text{P}\left(z < \frac{44 - 20}{\sqrt{341}}\right)$$

Using $z = \frac{x - \mu}{\sigma}$.

$$= \text{P}(z < 1.30)$$
$$= 0.9032$$

You do not interpolate so round your z value to 2 decimal places and use the tables, or use a graphical calculator.

b
$$\text{E}(X) = 20$$
$$\text{Var}(X) = \text{Var}(A) + 3\text{Var}(B) + 16\,\text{Var}(C)$$
$$= 2^2 + 3 \times 3^2 + 16 \times 4^2$$
$$C = 287$$

$$X = A - \sum_{i=1}^{3} B_i + 4C.$$
$$= A - (B_1 + B_2 + B_3) + 4C$$

$$\text{P}(X > 0) = \text{P}\left(z > \frac{-20}{\sqrt{287}}\right)$$

Using $z = \frac{x - \mu}{\sigma}$.

$$= \text{P}(z > -1.18)$$
$$= 0.8810$$
$$= 0.881 \text{ (3 s.f.)}$$

You do not interpolate so round your z value to 2 decimal places and use the tables, or use a graphical calculator.

Solutionbank S3
Edexcel AS and A Level Modular Mathematics

Review Exercise 1
Exercise A, Question 19

Question:

A manufacturer produces two flavours of soft drink, cola and lemonade. The weights, C and L, in grams, of randomly selected cola and lemonade cans are such that $C \sim N(350, 8)$ and $L \sim N(345, 17)$.

a Find the probability that the weights of two randomly selected cans of cola will differ by more than 6 g.

One can of each flavour is selected at random.

b Find the probability that the can of cola weighs more than the can of lemonade.

Cans are delivered to shops in boxes of 24 cans. The weights of empty boxes are normally distributed with mean 100 g and standard deviation 2 g.

c Find the probability that a full box of cola cans weighs between 8.51 kg and 8.52 kg.

d State an assumption you made in your calculation in part c. ***E***

Solution:

a Let $W = C_1 - C_2$

$E(W) = 350 - 350 = 0$

$Var(W) = 8 + 8$

$= 16$

$\therefore W \sim N(0, 16)$

Using $E(X - Y) = E(X) - E(Y)$

Using $Var(X - Y) = Var(X) + Var(Y)$

$\therefore P(|W| > 6) = 2P(W > 6)$

$= 2 \times P\left(Z > \frac{6-0}{\sqrt{16}}\right)$

$= 2 \times P(Z > 1.5)$

$= 2 \times (1 - 0.9332)$

$= 0.1336$

Using $z = \frac{x - \mu}{\sigma}$.

b Let $W = C - L$

$E(W) = 350 - 345 = 5$

$Var(W) = 8 + 17$

$= 25$

$\therefore W \sim N(5, 25)$

$P(W > 0) = P\left(Z > \frac{-5}{\sqrt{25}}\right)$

$= P(Z < -1)$

$= 0.8413$

Using $E(X - Y) = E(X) - E(Y)$

Using $Var(X - Y) = Var(X) + Var(Y)$

If you have used $W = L - C$ then

$P(W < 0) = P\left(Z < \frac{(0-(-5))}{\sqrt{25}}\right)$

$= P(Z < 1)$.

c Let $W = C_1 + \ldots + C_{24} + B$

A full box consists of 24 cans and a box.

$\therefore E(W) = 24 \times 350 + 100 = 8500$

$Var(W) = 24 \times 8 + 2^2 = 196$

Since we are adding we use $E(X_1 + X_2) = E(X_1) + E(X_2)$ and $Var(X_1 + X_2) = Var(X_1) + Var(X_2)$

$P(8510 \le W \le 8520) = P\left(\frac{8510 - 8500}{\sqrt{196}} \le Z \le \frac{8520 - 8500}{\sqrt{196}}\right)$

$= P(0.71\ldots \le Z \le 1.43\ldots)$

$= 0.9236 - 0.7611$

$= 0.1625$

(or 0.16096 directly from calculator)

Units must be consistent so change them all to grams and use $z = \frac{x - \mu}{\sigma}$.

You do not interpolate so round your z value to 2 decimal places and use the tables, or use a graphical calculator.

d All random variables are independent and normally distributed.