Pearson Edexcel

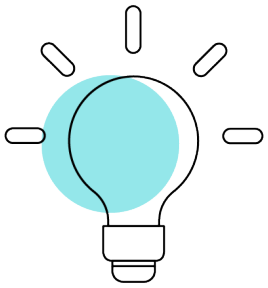
Topic Test Summer 2022

Pearson Edexcel GCE Mathematics (9MA0)

Paper 3 – Statistics

Topic 4: Discrete probability distributions; normal approximation

Contents



Pearson: helping people progress, everywhere

Pearson aspires to be the world's leading learning company. Our aim is to help everyone progress in their lives through education. We believe in every kind of learning, for all kinds of people, wherever they are in the world. We've been involved in education for over 150 years, and by working across 70 countries, in 100 languages, we have built an international reputation for our commitment to high standards and raising achievement through innovation in education. Find out more about how we can help you and your students at: www.pearson.com/uk

General guidance to Topic Tests

Context

- Topic Tests have come from past papers both published (those materials available publicly) and unpublished (those currently under padlock to our centres) presented in a different format to allow teachers to adapt them for use with candidates.

Purpose

- The purpose of this resource is to provide qualification-specific sets/groups of questions covering the knowledge, skills and understanding relevant to this Pearson qualification.
- This document should be used in conjunction with the advance information for the subject as well as general marking guidance for the qualification (available in published mark schemes).

Revise Revision Guide content coverage

The questions in this topic test have been taken from past papers, and have been selected as they cover the topic(s) most closely aligned to the A level advance information for summer 2022:

- Topic 4: Discrete probability distributions; normal approximation

The focus of content in this topic test can be found in the Revise Pearson Edexcel A level Mathematics Revision Guide. Free access to this Revise Guide is available for front of class use, to support your students' revision.

Contents	Revise Guide page reference	Level
Pure Mathematics	1-111	A level
Statistics	112-147	A level
Mechanics	148-181	A level

Content on other pages may also be useful, including for synoptic questions which bring together learning from across the specification.

Questions

Question T4_Q1

1. Helen believes that the random variable C, representing cloud cover from the large data set, can be modelled by a discrete uniform distribution.

(a) Write down the probability distribution for C.

(2)

(b) Using this model, find the probability that cloud cover is less than 50%

(1)

Helen used all the data from the large data set for Hurn in 2015 and found that the proportion of days with cloud cover of less than 50% was 0.315

(c) Comment on the suitability of Helen's model in the light of this information.

(1)

(d) Suggest an appropriate refinement to Helen's model.

(1)

Question 1 continued

Question T4_Q2

3. In an experiment a group of children each repeatedly throw a dart at a target.
For each child, the random variable H represents the number of times the dart hits the target in the first 10 throws.

Peta models H as B(10, 0.1)

(a) State two assumptions Peta needs to make to use her model.

(2)

(b) Using Peta's model, find P($H \geqslant 4$)

(1)

For each child the random variable F represents the number of the throw on which the dart first hits the target.

Using Peta's assumptions about this experiment,

(c) find P($F = 5$)

(2)

Thomas assumes that in this experiment no child will need more than 10 throws for the dart to hit the target for the first time. He models P($F = n$) as

$$\mathrm{P}(F = n) = 0.01 + (n - 1) \times \alpha$$

where α is a constant.

(d) Find the value of α

(4)

(e) Using Thomas' model, find P($F = 5$)

(1)

(f) Explain how Peta's and Thomas' models differ in describing the probability that a dart hits the target in this experiment.

(1)

Question 3 continued

Question 3 continued

Question 3 continued

Question T4_Q3

4. Magali is studying the mean total cloud cover, in oktas, for Leuchars in 1987 using data from the large data set. The daily mean total cloud cover for all 184 days from the large data set is summarised in the table below.

Daily mean total cloud cover (oktas)	0	1	2	3	4	5	6	7	8
Frequency (number of days)	0	1	4	7	10	30	52	52	28

One of the 184 days is selected at random.

(a) Find the probability that it has a daily mean total cloud cover of 6 or greater.

(1)

Magali is investigating whether the daily mean total cloud cover can be modelled using a binomial distribution.

She uses the random variable X to denote the daily mean total cloud cover and believes that $X \sim \text{B}(8, 0.76)$

Using Magali's model,

(b) (i) find $\text{P}(X \geqslant 6)$

(2)

(ii) find, to 1 decimal place, the expected number of days in a sample of 184 days with a daily mean total cloud cover of 7

(2)

(c) Explain whether or not your answers to part (b) support the use of Magali's model.

(1)

There were 28 days that had a daily mean total cloud cover of 8
For these 28 days the daily mean total cloud cover for the **following** day is shown in the table below.

Daily mean total cloud cover (oktas)	0	1	2	3	4	5	6	7	8
Frequency (number of days)	0	0	1	1	2	1	5	9	9

(d) Find the proportion of these days when the daily mean total cloud cover was 6 or greater.

(1)

(e) Comment on Magali's model in light of your answer to part (d).

(2)

Question 4 continued

Question 4 continued

Question 4 continued

Question T4_Q4

1. (a) State one disadvantage of using quota sampling compared with simple random sampling.

(1)

In a university 8% of students are members of the university dance club.

A random sample of 36 students is taken from the university.

The random variable X represents the number of these students who are members of the dance club.

(b) Using a suitable model for X, find

(i) $P(X = 4)$

(ii) $P(X \geqslant 7)$

(3)

Only 40% of the university dance club members can dance the tango.

(c) Find the probability that a student is a member of the university dance club and can dance the tango.

(1)

A random sample of 50 students is taken from the university.

(d) Find the probability that fewer than 3 of these students are members of the university dance club and can dance the tango.

(2)

Question 1 continued.

Question T4_Q5

6. The discrete random variable X has the following probability distribution

x	a	b	c
$P(X = x)$	$\log_{36} a$	$\log_{36} b$	$\log_{36} c$

where

- a, b and c are distinct integers $(a < b < c)$
- all the probabilities are greater than zero

(a) Find

(i) the value of a

(ii) the value of b

(iii) the value of c

Show your working clearly.

(5)

The independent random variables X_1 and X_2 each have the same distribution as X

(b) Find $P(X_1 = X_2)$

(2)

Question 6 continued.

Question 6 continued.

Mark Scheme

Question T4_Q1

Qu 1	Scheme	Marks	AO
(a)	c: 0 1 2 3 4 5 6 7 8 P($C = c$): $\frac{1}{9}$ $\frac{1}{9}$ $\frac{1}{9}$ $\frac{1}{9}$ $\frac{1}{9}$ $\frac{1}{9}$ $\frac{1}{9}$ $\frac{1}{9}$ $\frac{1}{9}$	B1 B1ft (2)	1.2 1.2
(b)	P($C < 4$) = $\frac{4}{9}$ (accept 0.444 or better)	B1 (1)	3.4
(c)	Probability lower than expected suggests model is <u>not</u> good	B1ft (1)	3.5a
(d)	e.g. Cloud cover will vary from month to month and place to place So e.g. use a non-uniform distribution	B1 (1)	3.5c
		(5 marks)	

Notes

(a) 1st B1 for a correct set of values for c. Allow $\{\frac{1}{8}, \frac{2}{8}, ... \frac{8}{8}\}$

2nd B1ft for correct probs from their values for c, consistent with discrete uniform distrib'n
Maybe as a prob. function. Allow P($X = x$) = $\frac{1}{9}$ for $0 \leqslant x \leqslant 8$ provided x = {0, 1, 2, …, 8} is clearly defined somewhere.

(b) B1 for using correct model to get $\frac{4}{9}$ (o.e.)

SC **Sample space {1, …, 8}** If scored B0B1 in (a) for this allow P($C < 4$) = $\frac{3}{8}$ to score B1 in (b)

(c) B1ft for comment that states that the model proposed is or is not a good one based on their model in part (a) and their probability in (b)
|(b) – 0.315| > 0.05 Allow e.g. "it is not suitable"; "it is not accurate" etc
|(b) – 0.315| ⩽ 0.05 Allow a comment that suggests it <u>is</u> suitable
No prob in (b) Allow a comparison that mentions 50% or 0.5 and rejects the model
No prob in (b) and no 50% or 0.5 or **(b) > 1** scores B0
Ignore any comments about location or weather patterns.

(d) B1 for a sensible refinement considering variations in month or location
Just saying "not uniform" is B0
Context & "non-uniform" Allow mention of different locations, months <u>and</u> non-uniform
<u>or</u> use more locations to form a new distribution with probabilities based on frequencies
Context & "binomial" Allow mention of different locations, months <u>and</u> binomial
Just refined model Model must be outlined and discrete and non-uniform
e.g. higher probabilities for more cloud cover <u>or</u> lower probabilities for less cloud cover
Continuous model Any model that is based on a continuous distribution. e.g. normal is B0

Question T4_Q2

Qu 3	Scheme	Marks	AO
(a)	The probability of a dart hitting the target is constant (from child to child and for each throw by each child) (o.e.) The throws of each of the darts are independent (o.e.)	B1 B1 (2)	1.2 1.2
(b)	[P($H \geqslant 4$) = 1 – P($H \leqslant 3$) = 1 – 0.9872 = 0.012795.. =] awrt **0.0128**	B1 (1)	1.1b
(c)	P($F = 5$) = $0.9^4 \times 0.1$, = 0.06561 = awrt **0.0656**	M1, A1 (2)	3.4 1.1b
(d)	<table><tr><td>n</td><td>1</td><td>2</td><td>…</td><td>10</td></tr><tr><td>P($F = n$)</td><td>0.01</td><td>0.01 + α</td><td>…</td><td>0.01+9α</td></tr></table> Sum of probs = 1 $\Rightarrow \frac{10}{2}[2\times0.01+9\alpha]=1$ [i.e. $5(0.02 + 9\alpha) = 1$ or $0.1 + 45\alpha = 1$] so α = **0.02**	M1 M1A1 A1 (4)	3.1b 3.1a 1.1b 1.1b
(e)	P($F = 5$| Thomas' model) = **0.09**	B1ft (1)	3.4
(f)	Peta's model assumes the probability of hitting target is constant (o.e.) **and** Thomas' model assumes this probability increases with each attempt(o.e.)	B1 (1)	3.5a
		(11 marks)	

Notes

(a) 1st B1 for stating that the probability (or possibility or chance) is constant (or fixed or same)
2nd B1 for stating that throws are independent ["trials" are independent is B0]

(b) B1 for awrt 0.0128 (found on calculator)

(c) M1 for a probability expression of the form $(1-p)^4 \times p$ where $0 < p < 1$
A1 for awrt 0.0656
SC Allow M1A0 for answer only of 0.066

(d) 1st M1 for setting up the distribution of F with at least 3 correct values of n and P($F = n$) in terms of α. (Can be implied by 2nd M1 or 1st A1)
2nd M1 for use of sum of probs = 1 **and** clear summation or use of arithmetic series formula (allow 1 error or missing term). (Can be implied by 1st A1)
1st A1 for a correct equation for α
2nd A1 for α = 0.02 (must be exact and come from correct working)

(e) B1ft for value resulting from $0.01 + 4 \times$ "their α" (provided α and the answer are probs)
Beware If their answer is the same as their (c) (or a rounded version of their (c)) score B0

(f) B1 for a suitable comment about the probability of hitting the target
ALT Allow idea that Peta's model suggests the dart may never hit the target but Thomas' says that it will hit at least once (in the first 10 throws).

Question T4_Q3

Question	Scheme	Marks	AOs
4 (a)	$\frac{132}{184} = 0.71739...$ awrt **0.717**	B1	1.1b
		(1)	
(b)(i)	$P(X \geq 6) = 1 - P(X \leq 5)$ or $P([X =]6) + P([X =]7) + P([X =]8)$	M1	3.4
	$= 1 - 0.296722...$ awrt **0.703**	A1	1.1b
		(2)	
(b)(ii)	$184 \times P(X = 7)$ $[= 184 \times 0.2811\ldots]$	M1	1.1b
	$= 51.7385...$ awrt **51.7**	A1	1.1b
		(2)	
(c)	Part (a) and part (b)(i) are similar **and** the expected number of 7s (51.7 or 0.281) matches with the number of 7s found in the data set (52 or 0.283) so Magali's model is supported.	B1ft	3.5a
		(1)	
(d)	$\frac{23}{28} = 0.82142...$ awrt **0.821**	B1	1.1b
		(1)	
(e)	Any one of… • Part (d)/'0.821' differs from part (a)/(b)(i)/(0.7…) • there is a **greater/different probability of high cloud cover/more likely to have high cloud cover** if the previous day had high cloud cover • **independence(o.e.)** does not hold	B1	2.4
	…therefore Magali's (binomial) model may not be suitable.	dB1	3.5a
		(2)	
			(9 marks)

Notes	
	Allow fractions, decimals or percentages throughout this question.
(a)	Allow equivalent fraction, e.g. $\frac{33}{46}$
(b)(i)	**M1:** for writing or using $1 - P(X \leq 5)$ or $P(X = 6) + P(X = 7) + P(X = 8)$ **A1:** awrt 0.703 (correct answer scores 2 out of 2)
(b)(ii)	**M1:** for $184 \times P(X = 7)$ o.e. e.g., $184 \times [P(X \leq 7) - P(X \leq 6)]$ **A1:** awrt 51.7
(c)	**B1ft:** comparing '0.717' with '0.703' and '51.7 or '0.281' with 52 or 0.283 **and** concluding that Magali's model **is supported** (must be comparing prob. with prob. and days with days). Allow not supported or mixed conclusions if consistent with their f.t. answers in (a) and (b)
(e)	**B1:** Any bullet point **dB1:** (dep on previous B1) for Magali's model may not be suitable (o.e.) Condone not accurate for not suitable **SC:** part (d) is similar to part (a)/(b)(i) **and** a compatible conclusion (i.e. Magali's model is supported) to score B1B1.

Question T4_Q4

Qu 1	Scheme	Marks	AO
(a)	**Disadvantage:** e.g. Not random; cannot use (reliably) for inferences	B1 **(1)**	1.1b
(b)	[Sight or correct use of] $X \sim B(36, 0.08)$	M1	3.3
(i)	$P(X = 4) = 0.167387\ldots$ awrt **0.167**	A1	1.1b
(ii)	$[P(X \ldots 7) = 1 - P(X \leq 6) =]$ 0.022233… awrt **0.0222**	A1 **(3)**	1.1b
(c)	P(In dance club and dance tango) = 0.4×0.08 = **0.032** or $\frac{4}{125}$ or **3.2%**	B1 **(1)**	1.1b
(d)	[Let T = those who can dance the Tango. Sight or use of] $T \sim B(50, \text{“}0.032\text{”})$	M1	3.3
	$[P(T < 3) = P(T \leq 2) =]$ 0.7850815… awrt **0.785**	A1 **(2)**	1.1b
		(7 marks)	

Notes

(a) B1 for a suitable disadvantage:

Allow (B1)	Do NOT allow (B0)
Not random or less random (o.e.)	Not representative
Cannot use (reliably) for inferences	Less accurate
(More likely to be) biased	Any comment based on time or cost
	Any mention of skew
	Any mention of non-response

(b) M1 for sight of B(36, 0.08) Allow in words: binomial with $n = 36$ and $p = 0.08$
may be implied by one correct answer to 2sf or sight of $P(X \leq 6) = 0.97776\ldots$ i.e. awrt 0.98
Allow for $36C4 \times 0.08^4 \times 0.92^{32}$ as this is "correct use"

(i) 1st A1 for awrt 0.167 NB An answer of just awrt 0.167 scores M1(⇒)1st A1

(ii) 2nd A1 for awrt 0.0222

(c) B1 for 0.032 o.e. (Can allow for sight of 0.4×0.08)

(d) M1 for sight of B(50, “0.032”) ft their answer to (c) provided it is a probability ≠ 0.08
may be implied by correct answer
or sight of $[P(T \leq 3)] = 0.924348\ldots$ i.e. awrt 0.924 or $P(T \leq 2)$ as part of $1 - P(T \leq 2)$ calc.
A1 for awrt 0.785

MR Allow MR of 50 (e.g. 30) provided clearly attempting $P(T \leq 2)$ and score M1A0

Question T4_Q5

Qu 6	Scheme	Marks	AO
(a)	[Sum of probs = 1 implies] $\log_{36} a + \log_{36} b + \log_{36} c = 1$	M1	3.1a
	$\Rightarrow \log_{36}(abc) = 1$ so $abc = 36$	A1	3.4
	All probabilities greater than 0 implies each of a, b and $c > 1$	B1	2.2a
	$36 = 2^2 \times 3^2$ (or 3 numbers that multiply to give 36 e.g. 2, 2, 9 etc)	dM1	2.1
	Since a, b and c are distinct must be **2, 3, 6** **($a = 2$, $b = 3$, $c = 6$)**	A1 **(5)**	3.2a
(b)	$(\log_{36} a)^2 + (\log_{36} b)^2 + (\log_{36} c)^2$	M1	3.4
	[= 0.0374137…+ 0.09398737…+0.25]		
	= 0.38140… awrt **0.381**	A1 **(2)**	1.1b
		(7 marks)	

	Notes
(a)	1st M1 for a start to the problem using sum of probabilities leading to eq'n in a, b and c
	1st A1 for reducing to the equation $abc = 36$ [Must follow from their equation.]
NB	Can go straight from $abc = 36$ to the answer for full marks for part (a).
	B1 for deducing that each value > 1 (may be implied by 3 integers all > 1 in the next line)
	2nd dM1 (dep on M1A1) for writing 36 as a product of prime factors or 3 values with product = 36 and none = 1
	2nd A1 for 2, 3 and 6 as a list or $a = 2$, $b = 3$ and $c = 6$
SC **Ans only**	**M0M0** If no method marks scored but a correct answer given score: M0A0B1M0A1 (2/5) This gets the SC score of 2/5 [Question says show your working clearly]
(b)	M1 for a correct expression in terms of a, b and c or values; ft their integers a, b and c Condone invisible brackets if the answer implies they are used.
	A1 for awrt 0.381