

Friday 22 June 2012 – Afternoon

A2 GCE MATHEMATICS

4733 Probability & Statistics 2

QUESTION PAPER

Candidates answer on the Printed Answer Book.

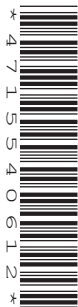
OCR supplied materials:

- Printed Answer Book 4733
- List of Formulae (MF1)

Other materials required:

- Scientific or graphical calculator

Duration: 1 hour 30 minutes



INSTRUCTIONS TO CANDIDATES

These instructions are the same on the Printed Answer Book and the Question Paper.

- The Question Paper will be found in the centre of the Printed Answer Book.
- Write your name, centre number and candidate number in the spaces provided on the Printed Answer Book. Please write clearly and in capital letters.
- **Write your answer to each question in the space provided in the Printed Answer Book.** Additional paper may be used if necessary but you must clearly show your candidate number, centre number and question number(s).
- Use black ink. HB pencil may be used for graphs and diagrams only.
- Answer **all** the questions.
- Read each question carefully. Make sure you know what you have to do before starting your answer.
- Do **not** write in the bar codes.
- You are permitted to use a scientific or graphical calculator in this paper.
- Give non-exact numerical answers correct to 3 significant figures unless a different degree of accuracy is specified in the question or is clearly appropriate.

INFORMATION FOR CANDIDATES

This information is the same on the Printed Answer Book and the Question Paper.

- The number of marks is given in brackets [] at the end of each question or part question on the Question Paper.
- **You are reminded of the need for clear presentation in your answers.**
- The total number of marks for this paper is **72**.
- The Printed Answer Book consists of **12** pages. The Question Paper consists of **4** pages. Any blank pages are indicated.

INSTRUCTION TO EXAMS OFFICER/INVIGILATOR

- Do not send this Question Paper for marking; it should be retained in the centre or recycled. Please contact OCR Copyright should you wish to re-use this document.

- 1 In one day's production, a machine produces 1000 CDs. Explain how to take a random sample of 15 CDs chosen from one day's production. [2]
- 2 (i) For the continuous random variable V , it is known that $E(V) = 72.0$. The mean of a random sample of 40 observations of V is denoted by \bar{V} . Given that $P(\bar{V} < 71.2) = 0.35$, estimate the value of $\text{Var}(V)$. [4]
- (ii) Explain why you need to use the Central Limit Theorem in part (i), and why its use is justified. [2]
- 3 It is known that on average one person in three prefers the colour of a certain object to be blue. In a psychological test, 12 randomly chosen people were seated in a room with blue walls, and asked to state independently which colour they preferred for the object. Seven of the 12 people said that they preferred blue. Carry out a significance test, at the 5% level, of whether the statement "on average one person in three prefers the colour of the object to be blue" is true for people who are seated in a room with blue walls. [7]
- 4 In a rock, small crystal formations occur at a constant average rate of 3.2 per cubic metre.
- (i) State a further assumption needed to model the number of crystal formations in a fixed volume of rock by a Poisson distribution. [1]
- In the remainder of the question, you should assume that a Poisson model is appropriate.
- (ii) Calculate the probability that in one cubic metre of rock there are exactly 5 crystal formations. [2]
- (iii) Calculate the probability that in 0.74 cubic metres of rock there are at least 3 crystal formations. [3]
- (iv) Use a suitable approximation to calculate the probability that in 10 cubic metres of rock there are at least 36 crystal formations. [5]
- 5 The acidity A (measured in pH) of soil of a particular type has a normal distribution. The pH values of a random sample of 80 soil samples from a certain region can be summarised as
- $$\Sigma a = 496, \quad \Sigma a^2 = 3126.$$
- Test, at the 10% significance level, whether in this region the mean pH of soil is 6.1. [11]
- 6 At a tourist car park, a survey is made of the regions from which cars come.
- (i) It is given that 40% of cars come from the London region. Use a suitable approximation to find the probability that, in a random sample of 32 cars, more than 17 come from the London region. Justify your approximation. [7]
- (ii) It is given that 1% of cars come from France. Use a suitable approximation to find the probability that, in a random sample of 90 cars, exactly 3 come from France. [4]

- 7 The continuous random variable X has probability density function

$$f(x) = \begin{cases} kx^2 & 0 \leq x \leq a, \\ 0 & \text{otherwise,} \end{cases}$$

where a and k are constants.

- (i) Sketch the graph of $y = f(x)$ and explain in non-technical language what this tells you about X . [3]
- (ii) Given that $E(X) = 4.5$, find
- (a) the value of a , [6]
- (b) $\text{Var}(X)$. [3]

- 8 The random variable X has the distribution $N(\mu, 8^2)$. A test is carried out, at the 5% significance level, of $H_0: \mu = 30$ against $H_1: \mu > 30$, based on a random sample of size 18.

- (i) Find the critical region for the test. [4]
- (ii) If $\mu = 30$ and the outcome of the test is that H_0 is rejected, state the type of error that is made. [1]

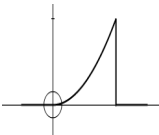
On a particular day this test is carried out independently a total of 20 times, and for 4 of these tests the outcome is that H_0 is rejected. It is known that the value of μ remains the same throughout these 20 tests.

- (iii) Find the probability that H_0 is rejected at least 4 times if $\mu = 30$. Hence state whether you think that $\mu = 30$, giving a reason. [3]
- (iv) Given that the probability of making an error of the type different from that stated in part (ii) is 0.4, calculate the actual value of μ , giving your answer correct to 4 significant figures. [4]

Question		Answer	Marks	Guidance	
1		Number CDs (sequentially) Select using random numbers	B1 B1 [2]	List needn't be stated, but must mention CDs. <i>Not</i> "select numbers randomly". Hat, etc: B1B0 Systematic: 66 or 67 B1, random start B1	Assume sequential unless stated otherwise. If "number CDs randomly", B1 max unless "sort by number". Stratified: apply scheme
2	(i)	$\left(\frac{71.2 - 72.0}{\sigma / \sqrt{40}} \right) = -0.3853$ [$\sigma = 13.13,$] $\text{Var}(V) = 172.4$	M1 A1 B1 A1 [4]	Standardise with Φ^{-1} & $\sqrt{40}$, allow cc, $\sqrt{\quad}$ errors eg σ^2 Square roots and sign correct, no cc, no "1 -" error z in range (\pm) [0.385, 0.386] seen <i>Final</i> answer in range [172, 173], or 13.1^2 cwo	RHS must be Φ^{-1} , i.e. <i>not</i> 0.7411 or 0.2589 or 0.6368 or 0.35. "1 -" error or $\times 40/39$: M1A0 [0.674 may be from "1 - 0.35 = 0.75"] Needs variance, not SD NB: Look out for $-13.1 \rightarrow 172$, M1A0B1A0
2	(ii)	Parent distribution not known n is large	B1 B1 [2]	Or clear equivalent. Not " <i>sample</i> not normal" Or clear equiv, e.g. sample size > 30 . Extras: max 1 " n large, $n > n_0$ ": B1 if $n_0 \geq 30$.	Don't bother about order of these statements. If numerical must be 30. Ignore "continuous".
3		α : $H_0: p = \frac{1}{3}$ [or 0.33 or better] $H_1: p \neq \frac{1}{3}$ [or 0.33 or better] B(12, $\frac{1}{3}$) stated or implied $P(\geq 7) = 1 - 0.9336 = 0.0664 > 0.025$	B2 M1 A1 A1	Allow π , but $\mu = \frac{1}{3}$ etc B1. Any other letter, B0 One-tailed, or no symbol, B1 max B(12, $\frac{1}{3}$) stated or implied, allow for N(4,8/3), Po(4) Probability in range [0.066, 0.067] Explicit comparison with 0.025, or $2p$ with 0.05	Not $\mu = 4$ (<i>if in doubt, consult</i>) <i>If N used, or $P(\leq 7)$ or $P(= 7)$, no more marks</i>
		β : CR is ≥ 8 , 7 not in CR Probability is 0.0188	A1 A1	Needs explicit comparison of 7 with CV Must be ≥ 7 , 0.019 or 0.0188 or better, allow 0.9812	1-tailed: A0 here regardless of value Need to be clear that CR is being used – look for comparison with 7. See also ζ .
		Do not reject H_0 . Insufficient evidence that statement is false.	M1 A1 [7]	Needs correct method, including like-with-like, correct tail, ≥ 7 (or ≤ 6). If CV, needs right tail A1 needs "evidence" or equivalent. "Statement" is enough context here	Allow from 1-tail. 0.9812 or 0.0188 or 0.0476: M0 unless " ≥ 7 " stated or clearly using β . $\sqrt{\quad}$ on their p /CR. Withhold if answer refers only to p .

Question		Answer	Marks	Guidance	
4	(i)	Crystals must occur independently of one another	B1 [1]	Allow interpreted, or “randomly” but nothing else. Must be contextualised; no other answers included.	Ignore “singly” (meaningless in this context). But allow “probability... is independent”
4	(ii)	$e^{-3.2} \frac{3.2^5}{5!} = \mathbf{0.114(0)}$	M1 A1 [2]	Formula, or .0608 or .1781 or .1075 or .1203 (tables) Answer a.r.t. 0.114, implies both marks	
4	(iii)	Po(2.368) $1 - e^{-2.368} \left(1 + 2.368 + \frac{2.368^2}{2}\right)$ $= \mathbf{0.4219}$	M1 M1 A1 [3]	Po(0.74 × 3.2) stated or implied 1 – correct Poisson terms, their λ , allow ± 1 term Answer, a.r.t. 0.422, implies all 3 marks	Allow for 0.75 × 3.2 etc, e.g. Po(2.4) Don’t allow second M1 from λ in tables, e.g. if MR, treat as E-1. If no working: don’t give M1A0
4	(iv)	Po(32) \approx N(32, 32) $1 - \Phi\left(\frac{35.5 - 32}{\sqrt{32}}\right)$ $= 1 - \Phi(.619) = \mathbf{0.2681}$	M1 A1 M1 A1 A1 [5]	N(λ , λ) stated or implied, allow $\sqrt{\lambda}$ or λ^2 for var N(32, 32), allow $\sqrt{32}$ or 32^2 for var Standardise with λ and $\sqrt{\lambda}$ or λ , allow cc errors but not \sqrt{n} ; both cc and $\sqrt{\quad}$ correct Final answer, a.r.t. 0.268	Needs $\lambda \geq 15$ Can get (M0A0) M1A1 from $\lambda < 15$ Typically, no cc \rightarrow 0.2203, or 32 \rightarrow 0.4565, 3/5 (but needs evidence, not just answer)

Question	Answer	Marks	Guidance
5	$H_0: \mu = 6.1$ $H_1: \mu \neq 6.1$ $\hat{\mu} = \bar{x} = 6.2$ $\hat{\sigma}^2 = \frac{80}{79} \left(\frac{3126}{80} - 6.2^2 \right) = 0.643$ $\alpha:$ $z = \frac{6.2 - 6.1}{\sqrt{0.643/80}} = 1.115$ $[1 - \Phi(1.115) = 0.1325 > 0.05]$ $1.115 < 1.645$	B2 B1 M1 M1 A1 M1 A1 A1	Both: B2. One error, B1, but \bar{x} , x , r etc: 0. 6.2: B0 6.2 [31/5] seen somewhere (other than hypotheses) Correct formula for biased estimate [0.635 or 127/200] Divide by 79 somewhere Variance estimate, a.r.t. 0.643, can be implied Standardise their 6.2 with reasonable variance attempt, needs 80, allow cc $z \in [1.11, 1.12]$ (not $-$) or $p \in [0.1323, 0.1333]$ Compare z with 1.645 (allow -1.645 if $z < 0$) or $p (< 0.5)$ with 0.05
	$\beta:$ CV $6.1 + 1.645 \times \sqrt{\frac{0.643}{80}}$ $= 6.247$ and $6.2 < 6.247$	M1 A1 A1√	$6.1 + z\sqrt{(\hat{\sigma}^2/80)}$, allow \pm , \sqrt errors CV, a.r.t. 6.25, needs $z = 1.645$, allow biased $\hat{\sigma}^2$ Compare 6.2 with CV from + sign, \sqrt on z (but not σ)
	Do not reject H_0 . Insufficient evidence that pH value is not 6.1	M1 A1√ [11]	Needs essentially correct method and comparison, needs 80 but no need for correct variance Needs context and “evidence” or equivalent, ft on their $z/p/CV$
Notes:	Biased estimate used : typically gets B2B1 M1M0A0 M1A0A1 M1A1 [total 8]	\bar{x} and μ interchanged: allow final M1A1 if <i>anywhere</i> right, but if always wrong (in hypotheses and z) M0A0. This would typically get B0B0B1 M1M1A1 M1A0A0 M0A0 [total 5]	

Question		Answer	Marks	Guidance	
6	(i)	$B(32, 0.4)$ $\approx N(12.8, 7.68)$ Valid as 12.8 and $19.2 > 5$ $1 - \Phi\left(\frac{17.5 - 12.8}{\sqrt{7.68}}\right)$ $[= 1 - \Phi(1.696)] = \mathbf{0.0449}$	B1 M1A1 B1 M1 A1 A1 [7]	$B(32, 0.4)$ stated or implied, e.g. by $Po(12.8)$ $N(\text{their attempt at } np, npq)$; $N(12.8, 7.68)$ Or “ n large and p close to 0.5 ”. Not npq or $7.68 > 5$. Standardise, their np, npq , allow wrong/no cc or no $\sqrt{\quad}$ 17.5 and \sqrt{npq} correct Answer, a.r.t. 0.045	Poisson $[0.09888]$, or exact $[0.046269]$: B1 max SC: $B(12.8, 7.68/32)$: M1A0 Allow np and nq both asserted > 5 $\div 32$: M0
6	(ii)	$B(90, 0.01)$ $\approx Po(0.9)$ $e^{-0.9} \frac{0.9^3}{3!} = \mathbf{0.0494}$	B1 M1 M1 A1 [4]	$B(90, 0.01)$ stated or implied. $Po(\text{their attempt at } np)$ Correct formula or use of tables, e.g. 0.1646 or 0.0112 Final answer in range $[0.049, 0.05]$ [i.e., <i>not</i> 0.05]	Exact $[0.049003]$: B1 max. Don't treat $p = 0.1$ as MR. If $np > 5$, M0M0 No working, wrong answer \Rightarrow M0A0, but right answer \Rightarrow M1A1 provided clearly SC: $B(90, 0.1)$, $N(9, 8.1)$, $[0.015, 0.016]$ cwo B2
7	(i)	 <p>Values of X close to a are more likely than those close to 0</p>	M1 A1 B1 [3]	Positive parabola (only), through 0 , nothing below x -axis Clear truncation at ends Withhold if concept misunderstood. Need to have probability of <i>values</i> (not of <i>occurring</i>); not just shape. Allow for U-shape but nothing else	$k < 0$: M0 even if $k > 0$ as well. Don't need any scales, vertical line at a etc. Can be vertical at A , needn't be horizontal at O . E.g.: “More likely to <i>occur</i> for x close to a ”: B0. Ignore extra comments like “exponential”
7	(ii)	(a) $\int_0^a kx^2 dx = 1 \Rightarrow k = \frac{3}{a^3}$ $\int_0^a \frac{3}{a^3} x^3 dx = \frac{9}{2} \Rightarrow a = 6$	M1 A1 M1 A1 A1 A1 [6]	Attempt to integrate kx^2 , ignore limits Correct limits and equate to 1 Attempt to integrate kx^3 , ignore limits Correct limits and equate to 4.5 One correct equation connecting k and a , can be implied Correctly obtain $a = 6$ only	Must attempt integration Must attempt integration Don't need k in terms of a here $ka^3 = 3$ or $ka^4 = 18$, a.e. simplified form No marks explicitly for $k [= 1/72$ or $0.01388\dots]$

Question		Answer	Marks	Guidance
7	(ii) (b)	$\int_0^6 \frac{1}{72} x^4 dx \quad [= \frac{108}{5}]$ $21.6 - 4.5^2 = \mathbf{1.35}$	M1 A1 A1 [3]	Attempt to integrate kx^4 , their a, k , can be algebraic Subtract 4.5^2 (given in question) 1.35 or exact equivalent only Must attempt integration; limits 0, $a\sqrt{\quad}$ Somewhere [=27/20]
8	(i)	$30 + 1.645 \times \frac{8}{\sqrt{18}} = 33.102$ so CR is $\bar{X} > 33.1$	M1 A1 A1 A1√ [4]	$30 + z \times 8/\sqrt{18}$, allow $\sqrt{\quad}$ errors, cc 1.645, requires + only 33.1 a.r.t. 33.10 \geq their RH CV√, allow \leq their LH CV <i>as well</i> , allow $>$, allow no letter or X but no other letter Allow \pm but not $-$ only. No 18: 0 in this part. Don't allow "accept if ≤ 33.1 , reject if > 33.1 " Inequality required in final line
	(ii)	Type I [error]	B1 [1]	Nothing else unless it's just an amplification. Allow "Type 1"
	(iii)	B(20, 0.05): $P(\geq 4) = 0.0159$ so unlikely that $\mu = 30$	M1 A1 A1√ [3]	B(20, 0.05) stated or implied. Not B(20, 1/5) Probability, a.r.t. 0.016 Justified conclusion, e.g. "I think $\mu = 30$ as not less than 0.01". FT on their p . No reason: A0. Not over-assertive. But "I think $\mu = 30$ as probability is small" is A0.
	(iv)	$\frac{33.1 - \mu}{8/\sqrt{18}} = -0.253$ $\mu = \mathbf{33.58}$	M1 A1 A1 A1 [4]	Needs Φ^{-1} , their CV, SD right or same as in (i), allow cc Signs correct, can be implied by answer $>$ their CV z in range $(\pm)[0.25, 0.26]$ Final answer $33.55 \leq \mu \leq 33.60$, 4 SF needed. Not 30. Allow omission of $\sqrt{18}$ only if omitted in (i). "1 -" errors: can get M1A0A1 Typically 32.62 probably gets 2/4.