

**ADVANCED GCE  
MATHEMATICS**

Probability & Statistics 2

**4733**

**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

**Wednesday 22 June 2011  
Morning**

**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. Pencil may be used for graphs and diagrams only.
- Read each question carefully. Make sure you know what you have to do before starting your answer.
- Answer **all** the questions.
- 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 destroyed.

- 1 In Fisher Avenue there are 263 houses, numbered 1 to 263. Explain how to obtain a random sample of 20 of these houses. [3]

- 2 The random variable  $Y$  has the distribution  $N(\mu, \sigma^2)$ . It is given that

$$P(Y < 48.0) = P(Y > 57.0) = 0.0668.$$

Find the value  $y_0$  such that  $P(Y > y_0) = 0.05$ . [7]

- 3 The random variable  $X$  has the distribution  $N(\mu, 5^2)$ . A hypothesis test is carried out of  $H_0: \mu = 20.0$  against  $H_1: \mu < 20.0$ , at the 1% level of significance, based on the mean of a sample of size 16. Given that in fact  $\mu = 15.0$ , find the probability that the test results in a Type II error. [7]

- 4 A continuous random variable  $X$  has probability density function

$$f(x) = \begin{cases} \frac{3}{16}(x-2)^2 & 0 \leq x \leq 4, \\ 0 & \text{otherwise.} \end{cases}$$

(i) Sketch the graph of  $y = f(x)$ . [2]

(ii) Calculate the variance of  $X$ . [5]

(iii) A student writes “ $X$  is more likely to occur when  $x$  takes values further away from 2”. Explain whether you agree with this statement. [1]

- 5 A travel company finds from its records that 40% of its customers book with travel agents. The company redesigns its website, and then carries out a survey of 10 randomly chosen customers. The result of the survey is that 1 of these customers booked with a travel agent.

(i) Test at the 5% significance level whether the percentage of customers who book with travel agents has decreased. [7]

(ii) The managing director says that “Our redesigned website has resulted in a decrease in the percentage of our customers who book with travel agents.” Comment on this statement. [1]

- 6 Records show that before the year 1990 the maximum daily temperature  $T^\circ\text{C}$  at a seaside resort in August can be modelled by a distribution with mean 24.3. The maximum temperatures of a random sample of 50 August days since 1990 can be summarised by

$$n = 50, \quad \Sigma t = 1314.0, \quad \Sigma t^2 = 36\,602.17.$$

(i) Test, at the 1% significance level, whether there is evidence of a change in the mean maximum daily temperature in August since 1990. [11]

(ii) Give a reason why it is possible to use the Central Limit Theorem in your test. [1]

7 The number of customer complaints received by a company per day is denoted by  $X$ . Assume that  $X$  has the distribution  $Po(2.2)$ .

(i) In a week of 5 working days, the probability there are at least  $n$  customer complaints is 0.146 correct to 3 significant figures. Use tables to find the value of  $n$ . [3]

(ii) Use a suitable approximation to find the probability that in a period of 20 working days there are fewer than 38 customer complaints. [5]

A week of 5 working days in which at least  $n$  customer complaints are received, where  $n$  is the value found in part (i), is called a 'bad' week.

(iii) Use a suitable approximation to find the probability that, in 40 randomly chosen weeks, more than 7 are bad. [6]

8 (a) A group of students is discussing the conditions that are needed if a Poisson distribution is to be a good model for the number of telephone calls received by a fire brigade on a working day.

(i) Alice says "Events must be independent". Explain why this condition may not hold in this context. [1]

(ii) State a different condition that is needed. Explain whether it is likely to hold in this context. [2]

(b) The random variables  $R$ ,  $S$  and  $T$  have independent Poisson distributions with means  $\lambda$ ,  $\mu$  and  $\lambda + \mu$  respectively.


(i) In the case  $\lambda = 2.74$ , find  $P(R > 2)$ . [3]

(ii) In the case  $\lambda = 2$  and  $\mu = 3$ , find  $P(R = 0 \text{ and } S = 1) + P(R = 1 \text{ and } S = 0)$ . Give your answer correct to 4 decimal places. [3]

(iii) In the general case, show algebraically that

$$P(R = 0 \text{ and } S = 1) + P(R = 1 \text{ and } S = 0) = P(T = 1). \quad [4]$$

1		Number all the houses sequentially, or use house numbers Select using random numbers  Ignore numbers > 263	B1 B1 B1 <b>3</b>	Any mention of using house numbers, or houses, or other numbering. (List can be implied). <i>Not</i> random numbering unless correct subsequent method (e.g. sort them numerically) Mention random numbers. <i>Not</i> “select numbers randomly”. Must be random method. NB: Using $263 \times$ calculator Rand # is biased: B0. But “Ran#(263)” is unbiased. Deal with problem of > 263, <i>or</i> repeats. “Select 20 random numbers between 1 and 263”: B1B0 [If this, need to mention repeats to get last B1] Example: “put numbers/house names (etc) into hat and select”: B1B0B0
2	$\alpha$	$\mu = \frac{48+57}{2} = 52.5$ $\Phi^{-1}(0.9332) = 1.5$ $4.5 \div 1.5 \quad [\sigma = 3]$	M1 A1 B1 M1	Use symmetry to find $\mu$ Obtain $\mu = 52.5$ 1.5 seen, e.g. in $4.5 \div 1.5$ $4.5 \div$ their $\Phi^{-1}$ , or $1.645 \div$ their $\Phi^{-1}$ , must be +ve, allow cc
	$\beta$	$\frac{57-\mu}{\sigma} = 1.5, \frac{48-\mu}{\sigma} = -1.5$ Solve simultaneously: $\mu = 52.5 \quad [\sigma = 3]$	M1 A1 B1 A1	$\frac{57-\mu}{\sigma} = z, \frac{48-\mu}{\sigma} = -z$ M1 for one, ignoring cc, $\sigma^2$ , sign or “1 -” errors, RHS must be $\Phi^{-1}$ ( <i>not</i> $\Phi$ ) [e.g. 0.8246 or 0.5267] or 0.0668 or 0.9332); A1 for both completely correct except for value of z. $z = 1.5$ or $-1.5$ in at least one equation Solve without obvious errors, get $\mu = 52.5$ , OK from wrong z [NB: 52.5 from both signs wrong: A0]
		$\mu + \frac{4.5}{1.5} \times 1.645$ $= 57.4(35)$	M1 B1 A1 <b>7</b>	$\mu + z\sigma$ [Their $\mu$ and $\sigma$ , anything recognisable as z] [expect to see $52.5 + 3 \times 1.645$ ] $z = 1.645$ seen Answer in range [57.4, 57.45], cwo
3		CV $20 - \frac{5}{\sqrt{16}} \times 2.326 = 17.0925$  $P(X > 17.0925)$ $= \Phi\left(\frac{17.0925 - 15}{5/\sqrt{16}}\right) = \Phi(1.674)$ Answer <b>0.0471</b>	M1 B1 A1 M1*  A1 dep M1 A1 <b>7</b>	Attempt $20 - 5z/\sqrt{16}$ , allow SD $\leftrightarrow$ var errors, allow $20 \pm 5z/\sqrt{16}$ , <i>not</i> $20 + 5z/\sqrt{16}$ , allow cc 2.326 seen CV a.r.t. 17.1 [NB: <i>not</i> 17.9075] Standardise any attempt at a CV (from $\mu = 20$ ) with 15 and any SD that would have got first M1, allow cc $z = 1.674$ seen or implied, e.g. by $p = 0.047$ or 0.953 or 0.9535, allow anything in range [1.67, 1.68] Probability < 0.5, or > 0.5 if their CV is < 15 Answer, a.r.t. 0.047 [including 0.0465 from CV 17.1] Notes: 16 missing: can get M0B1A0M1A0M1A0, or even last two A1’s if 16 used then

4	(i)		M1 A1	2	Positive parabola, all above axis. [Don't worry about being pointed unless extreme.] Correct place, touches x-axis, not beyond the limits suggested by their axes, symmetric ends, not too straight
	(ii)	$\frac{3}{16} \int_0^4 x^2(x-2)^2 dx$ $= \frac{3}{16} \left[ \frac{x^5}{5} - x^4 + 4 \frac{x^3}{3} \right]_0^4 \quad [= 6\frac{2}{5}]$ $\sigma^2 = 6\frac{2}{5} - 2^2$ $= 2\frac{2}{5}$	M1 M1 B1 B1 A1	5	Attempt $\int x^2 f(x) dx$ , limits 0 and 4 Method for integration, e.g. multiply out [ <i>indept</i> ] [Or use $\sigma^2 = \frac{3}{16} \int_0^4 (x-2)^4 dx$ ] Correct indefinite integral, limits not needed, e.g. parts: $\frac{3}{16} \left[ \frac{x^2(x-2)^3}{3} - \frac{x(x-2)^4}{6} + \frac{(x-2)^5}{30} \right]$ Subtract $2^2$ Final answer 2.4, any equivalent exact form, cwo
	(iii)	No because $x$ represents a value taken by the random variable [ <i>not an event that "occurs"</i> ]	B1	1	Show clear understanding that $x$ is a value of $X$ . Usual misunderstanding is " $X$ is an event that may or may not occur, depending on $x$ ". However: SR: Allow B1 for answer clearly indicating that probabilities higher where curve higher, or clearly stating that all probabilities are effectively zero. E.g.: "Agree as area under graph [or " $f(x)$ "] increases", or "minimum at 2" B1 "True only between 0 and 4": B0 unless explanation Mention of variance etc: 0. "Agree because the graph shows this": B0
5	(i)	$H_0: p = 0.4; H_1: p < 0.4$ $B(10, 0.4)$	B1B1 M1		Both: B2. Allow $\pi$ . One error, B1, but $x$ or $r$ : 0. <b>SEE NOTES AT START AND END</b> $B(10, 0.4)$ stated or implied, e.g. $N(4, 2.4)$ [ $P(=1)$ [=0.0404] or $P(\geq 1)$ [=0.9940] or $P(<1)$ [=0.0060] or Poisson or normal, or RH tail for CR, gets no more marks in (i)]
	$\alpha$	$P(\leq 1) = 0.0464$ $< 0.05$ so reject $H_0$	A1 A1		This probability or 0.9536 only Explicit comparison with 0.05, or 0.9536 with 0.95
	$\beta$	CR is $\leq 1$ and compare 1 Probability of this is 0.0464	A1 A1		Comparison needn't be explicit in this method This probability needs to be seen
		Reject $H_0$ . Significant evidence that % who book with travel agents reduced	M1 A1✓	7	Correct method, ✓, comparison and first conclusion Interpreted in context, "evidence that" or equiv needed, ✓ on numbers
(ii)	Can't deduce cause-and-effect	B1	1	Equivalent comment, regardless of answer to (i). Ignore wrong answer if right answer seen "Other factors haven't been considered" B1 "Sample is small", or "test may be wrong" B0	

6	(i)	$H_0: \mu = 24.3; H_1: \mu \neq 24.3$ $\bar{t} = 26.28$ $\hat{\sigma}^2 = \frac{50}{49} \left[ \frac{36602.17}{50} - 26.28^2 \right]$ $= 42.25$ $z = \frac{26.28 - 24.3}{\sqrt{42.25/50}} = 2.154$ $< 2.576$	B1B1 B1 M1 M1  A1 M1 A1 A1	Both: B2. 1 error, B1, but $t, x$ etc: B0 26.28 seen or implied Correct formula for biased estimate [= 41.405] Multiply by 50/49 [Single formula: M2, or give M1 if wrong but 49 divisor seen] 42.25 or 6.5 seen or implied Standardise their $\bar{t}$ with 24.3, $\sqrt{50}$ , allow sign/ $\sqrt{\text{cc}}$ errors, their variance 2.15(4) or $p$ in range [0.0153, 0.0158], <i>not</i> -2.154 unless 0.015(6) subsequently used, <i>not</i> 1-tail Compare $z$ with $\pm 2.576$ , or $p > 0.005$ , or $2p$ with 0.01, <i>not</i> from $\mu = 26.28$	<b>SEE NOTES AT START AND END</b>
	$\beta$	CV $24.3 + 2.576 \times \sqrt{\frac{42.25}{50}}$ $= 26.67$ and $26.28 < 26.67$	M1 A1 A1	$24.3 + zs/\sqrt{50}$ , allow cc, $\sqrt{\text{errors}}$ , allow $\pm$ but not $-$ only. <i>Not</i> $26.28 - zs/\sqrt{50}$ $z = 2.576$ , <i>not</i> from $\mu = 26.28$ or 50 omitted, <i>not</i> from 1-tail Correct CV, $\checkmark$ on $z$ , and compare sample mean	
		Do not reject $H_0$ . Insufficient evidence of a change in maximum daily temperature.	M1 A1 $\checkmark$	Conclusion, $\checkmark$ , needs method, like-with-like, 50, <i>not</i> from $\mu = 26.28$ , <i>doesn't</i> need correct $z$ Contextualised, recognise uncertainty, $\checkmark$ on numbers NB: Clear evidence of $\mu = 26.28$ : can't get last 4 marks. <i>See exemplars <math>\gamma</math> and <math>\delta</math></i>	<b>11</b>
	(ii)	$n$ is large	B1	This answer <i>only</i> or " $n > \text{number}$ " where number $\geq 29$ , <i>not</i> both this and "distribution unknown". But " $n$ is large so we can approximate even though we don't know the distribution" is B1 "Possible as $n = 50$ " B0.	<b>1</b>
7	(i)	Po(11) $1 - P(\leq r) = 0.854$ gives $r = 14$ so $n = 15$	M1 M1 A1	Po(11) stated or clearly implied Find $1 - 0.146$ in tables, e.g. answer 14 [RH tail, e.g. "7", or single value only: max M1M0A0] $n = 15$ only, allow " $\geq 15$ "	<b>3</b>
	(ii)	Po(44) $\approx$ N(44, 44) $\Phi\left(\frac{37.5 - 44}{\sqrt{44}}\right) = \Phi(-0.980)$  $= 0.1635$	M1 A1 M1 A1 A1	Normal, mean attempted $2.2 \times 20$ Both parameters 44, allow var = $\sqrt{44}$ or $44^2$ Standardise, their 44, allow cc, $\sqrt{\text{errors}}$ , e.g. ans 0.283 or 0.2036 or 0.4411, <i>not</i> $\div 20$ $\sqrt{\text{and cc}}$ both correct Answer in range [0.163, 0.164]	<b>5</b>
	(iii)	B(40, 0.146) $\approx$ N(5.84, 4.98736) $1 - \Phi\left(\frac{7.5 - 5.84}{\sqrt{4.98736}}\right) = 1 - \Phi(0.7433)$  $= 0.2286$	M1 M1 A1 M1 A1 A1	B(40, 0.146) stated or implied, e.g. by Po(5.84) Normal, attempt at mean = $np$ [Poisson etc, or exact binomial (0.22132): no more marks] Both parameters correct [Poisson(5.84) $\rightarrow$ N(5.84, 5.84): M0A0] Standardise with their $np$ and $npq$ , allow cc, $\sqrt{\text{errors}}$ , e.g. ans 0.3838 or 0.302 or 0.370 $\text{cc and } \sqrt{\text{both}}$ correct Answer in range [0.228, 0.229] SC: B(40, 0.854) $\approx$ N(34.16, 4.98736): can get full marks, but if $R > 7$ used, max 3	<b>6</b>

8	(a) (i)	Several calls may all refer to the same incident	B1	1	Any reason showing correct understanding of “independent”, but not just “singly” or equivalent. Ignore extra condition(s) unless clearly wrong in which case B0. Not “fires” independent. “Fires might spread” B0
	(ii)	Calls occur at constant average rate	B1		This condition only, allow “average” omitted, <i>not</i> “constant probability”, <i>not</i> “random” unless clearly correct interpretation follows. No third condition unless fully justified by subsequent answer. Need contextualising <i>somewhere</i> in this part.
		E.g. No, because incidents are less/more common at night	B1	2	Any comment (with either yes or no) showing correct understanding, but “Fires might not occur at constant average rate” is not enough (gets B1 B0) “Different rates at different times of year”: B0
	(b) (i)	$1 - \left( 1 + 2.74 + \frac{2.74^2}{2!} \right) e^{-2.74}$ $= \mathbf{0.516(1)}$	M1 M1		Formula for any one correct Poisson probability for $r \geq 1$ [1 – (0.06457 + 0.17692 + 0.24238)] Correct overall formula, allow 1 error (e.g. 1 term extra or missing or no “1 –”)
	(ii)	$(e^{-2} \times 1)(e^{-3} \times 3) + (e^{-2} \times 2)(e^{-3} \times 1)$ $= \mathbf{0.0337}$	M1 A1 A1	3	Answer, a.r.t. 0.516 [Interpolation (0.51604) or no working: B0 or B3]
	(iii)	$(e^{-\lambda} \times 1)(e^{-\mu} \times \mu) + (e^{-\lambda} \times \lambda)(e^{-\mu} \times 1)$ $= e^{-\lambda} \times e^{-\mu} (\lambda + \mu)$ $= e^{-(\lambda + \mu)} (\lambda + \mu)$ $= P(T = 1)$	M1 M1 A1 A1	4	Correct algebraic expression [Ignore 1! throughout] Take out factor of $e^{-\lambda} \times e^{-\mu}$ or equivalent essential step Correctly obtain exact answer [allow $e^{-\lambda - \mu}(\lambda + \mu)$ ] All correct, and write down correct formula for $P(T = 1)$ [NB: $T$ needed] Allow working towards middle SR: $\lambda = 2, \mu = 3$ : Can get M1M1A1A0 if $e^{-2}$ and $e^{-3}$ retained. As soon as decimal approximations seen, no more marks.

*Specific examples for question 5(i)*

$\alpha$	$H_0: p = 0.4; H_1: p < 0.4$ $N(4, 2.4)$ $P(\leq 1) = 0.0533$ $> 0.05$ So do not reject $H_0$ . Insufficient evidence that % who book with travel agents reduced	B1B1 M1 A0  M0  <b>3</b>	$\delta$	$H_0: p = 0.4; H_1: p < 0.4$ $B(10, 0.4)$ $P(\geq 1) = 0.9939$ $> 0.95$ So reject $H_0$ Insufficient evidence that % who book with travel agents reduced	B1B1 M1 A0 A0 M0 A0  <b>3</b>
$\beta$	$H_0: p = 0.4; H_1: p < 0.4$ $B(10, 0.4)$ “ $P(= 1) = 0.0464$ ” <i>[allow this]</i> $< 0.05$ So reject $H_0$ Insufficient evidence that % who book with travel agents reduced	B1B1 M1 A1 A1 M1 A0  <b>6</b>	$\epsilon$	$H_0: p = 0.4; H_1: p \neq 0.4$ <i>[two-tailed]</i> $B(10, 0.4)$ “ $P(= 1) = 0.0464$ ” $> 0.025$ So do not reject $H_0$ Insufficient evidence that % who book with travel agents reduced	B1B0 M1 A1 A0 M1 A1  <b>5</b>
$\gamma$	$H_0: p = 0.4; H_1: p < 0.4$ $B(10, 0.4)$ $P(= 1) = 0.0404$ <i>[look out for this]</i> $< 0.05$ so reject $H_0$ Significant evidence that % who book with travel agents reduced	B1B1 M1 A0 A0 M0 A0  <b>3</b>	$\zeta$	$H_0: p = 0.4; H_1: p < 0.4$ $B(10, 0.4)$ $P(= 1) = 0.0464$  <i>[no explicit comparison]</i> So reject $H_0$ . Significant evidence that % who book with travel agents reduced	B1B1 M1 A1 A0 M1 A1  <b>6</b>



Specific examples for question 6(i)

<p><b>α</b></p>	<p><math>H_0: \bar{t} = 24.3; H_1: \bar{t} \neq 24.3</math> [wrong symbol]  <math>\bar{t}</math> not seen explicitly [implied by ...]  <math>\hat{\sigma}^2 = \left[ \frac{36602.17}{50} - 26.28^2 \right] = 41.405</math> [biased est]  <math>z = \frac{26.28 - 24.3}{\sqrt{41.405/50}} = 2.1758</math>  <math>&lt; 2.576</math>                      Accept <math>H_0</math>, maximum temp unchanged                      [over-assertive, otherwise A1]</p>	<p>B0B0                      B1                      M1                      M0                      A0                      M1                      A0                      A1                      M1A0 <b>5</b></p>	<p><b>δ</b></p>	<p><math>H_0 = 24.3; H_1 \neq 24.3</math> [missing symbol]  <math>\bar{t} = 26.28</math>  <math>\hat{\sigma}^2 = \dots = 42.25</math>  <math>z = \frac{24.3 - 26.28}{\sqrt{42.25/50}} = -2.154</math> [loses 1]  <math>&gt; -2.576</math>                      Insufficient evidence to reject <math>H_0</math>. No change in maximum daily temperature. [OK]</p>	<p>B1 only                      B1                      M1M1                      A1                      M1                      A0                      A1                      M1                      A1 <b>9</b></p>
<p><b>β</b></p>	<p><math>H_0: \mu = 26.28; H_1: \mu \neq 26.28</math> [WRONG]  <math>\bar{t} = 24.3</math> [explicitly]  <math>\hat{\sigma}^2 = \dots = 42.25</math>  <math>z = \frac{26.28 - 24.3}{\sqrt{42.25/50}} = 2.154</math> [allow this – BOD]  <math>&lt; 2.576</math>                      Accept <math>H_0</math>. Insufficient evidence of a change in maximum daily temperature.</p>	<p>B0B0                      B0                      M1M1                      A1                      M1                      A1                      A1                      M1                      A1 <b>8</b></p>	<p><b>ε</b></p>	<p><math>H_0: \mu = 24.3; H_1: \mu &gt; 24.3</math> [one-tail]  <math>\bar{t} = 26.28</math>  <math>\hat{\sigma}^2 = \dots = 42.25</math>  <math>z = \frac{26.28 - 24.3}{\sqrt{42.25/50}} = 2.154</math>  <math>&lt; 2.326</math>                      Accept <math>H_0</math>. Insufficient evidence of a change in maximum daily temperature.</p>	<p>B1B0                      B1                      M1M1                      A1                      M1                      A1                      A0                      M1                      A1 <b>9</b></p>
<p><b>γ</b></p>	<p><math>H_0: \mu = 26.28; H_1: \mu \neq 26.28</math> [WRONG]  <math>\bar{t}</math> not seen separately [implied]  <math>\hat{\sigma}^2 = \dots = 42.25</math>  <math>z = \frac{24.3 - 26.28}{\sqrt{42.25/50}} = -2.154</math> [DON'T allow this]  <math>&gt; -2.576</math>                      Accept <math>H_0</math>. Insufficient evidence of a change in maximum daily temperature.</p>	<p>B0B0                      B1                      M1M1                      A1                      M1                      A0                      A0                      M0                      A0 <b>5</b></p>	<p><b>ζ</b></p>	<p><math>z = \frac{24.3 - 26.28}{\sqrt{42.25/50}} = -2.154</math> but then...                      So <math>p = 0.0156 &gt; 0.005</math> [OK here]                      Accept <math>H_0</math>. Insufficient evidence of a change in maximum daily temperature.</p>	<p>M1                      A1                      A1                      M1                      A1 <b>(11)</b></p>
			<p><b>η</b></p>	<p><math>z = \frac{26.28 - 24.3}{\sqrt{42.25}} = 0.3046</math> [no <math>\sqrt{50}</math>]  <math>&lt; 2.576</math>                      Accept <math>H_0</math>. Insufficient evidence of a change in maximum daily temperature.</p>	<p>M0                      A0                      A0                      M0                      A0 <b>(6)</b></p>

The following guidance notes are provided.

**1 Standardisation using the normal distribution.**

- (a) In *stating* parameters of normal distributions, don't worry about the difference between  $\sigma$  and  $\sigma^2$ , so allow  $N(9, 16)$  or  $N(9, 4^2)$  or  $N(9, 4)$ . When *calculating*  $\frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$ , the following mistakes are accuracy mistakes and not method mistakes so can generally score M1A0:  
confusion of  $\sigma$  with  $\sigma^2$  or  $\sqrt{\sigma}$ ;  $n$  versus  $\sqrt{n}$ ; wrong or no continuity corrections.
- (b) Use of  $\frac{\mu - \bar{x}}{\sigma}$  instead of  $\frac{\bar{x} - \mu}{\sigma}$  is not penalised if it leads to a correct probability, but if the candidate is using a  $z$ -value in a hypothesis test, an answer of  $z = -2.15$  when it ought to be 2.15 is an accuracy error and loses the relevant A1. When finding  $\mu$  or  $\sigma^2$  from probabilities, some candidates are taught to use  $\frac{\mu - \bar{x}}{\sigma}$  whenever  $\mu > \bar{x}$ ; provided the signs are consistent this gains full marks.
- (c) Some candidates are taught to calculate, for example,  $P(X > 5)$  from  $N(9, 16)$  by calculating instead  $P(X < 13)$ . This is a correct method, though it looks very strange the first time you see it.
- (d) When calculating normal approximations to binomial or Poisson, use of the wrong, or no, continuity correction generally loses the last two marks: A0 A0.

**2 Conclusions to hypothesis tests.** There are generally 2 marks for these.

- (a) In order to gain M1, candidates must not only say the correct "Reject/do not reject  $H_0$ " but have done the whole test in essence correctly apart from numerical errors. In other words, they must have compared their  $p$  value with a critical  $p$  value or other "like-with-like" (e.g. *not* say 0.0234 with 1.96), using the correct tail (e.g. *not*  $-2.61$  with  $+2.576$ ), and the working should in general have accuracy errors only. Thus miscalculation of  $z$ , comparison with 1.645 instead of 1.96, or using  $n$  instead of  $\sqrt{n}$ , or omission of a continuity correction when it is necessary, are all accuracy errors and the candidate can still gain the last M1 A1. Omission of  $\sqrt{n}$  where it is necessary is a method mistake and so gets M0. In hypothesis tests using discrete distributions, use of  $P(\leq 12)$  or  $P(> 12)$  or  $P(= 12)$  when it should be  $P(\geq 12)$  is a method mistake and usually loses all the final marks in a question.
- (b) The A1 mark is for interpreting the answer *in the context of the question*, and *without over-assertiveness*. Thus "The mean number of applicants has increased" is over-assertive and gets A0 (although we allow "There is sufficient evidence to reject  $H_0$ . The mean number of applicants has increased", A1), and "There is sufficient evidence that the mean has increased" is not contextualised, so that too is A0.
- (c) A wrong statement such as  $-2.61 > -2.576$  generally gets B0 for comparison but can get the subsequent M1A1. Otherwise:
- (d) If there is a self-contradiction, award M1 only if "Reject/Accept  $H_0$ " is consistent with their comparison. Thus if, say, we had  $z = 2.61 > z_{\text{crit}} = 2.576$ :  
"Reject  $H_0$ , there is insufficient evidence that the mean number of ... has changed" is M1A0.  
but "Do not reject  $H_0$ , there is evidence that the mean number of ... has changed" is M0A0.
- (e) We don't usually worry about differences between "Reject  $H_0$ " and "Accept  $H_1$ " etc.