Oxford, Cambridge and RSA Examinations

Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education

# MEI STRUCTURED MATHEMATICS 2615

Statistics 3

Tuesday 19 JUNE 2001 Morning 1 hour 20 minutes

Additional materials:
Answer booklet
Graph paper
MEI Examination Formulae and Tables (MF12)

**TIME** 1 hour 20 minutes

## INSTRUCTIONS TO CANDIDATES

Write your Name, Centre Number and Candidate Number in the spaces provided on the answer booklet.

Answer **all** questions.

You are permitted to use a graphical calculator in this paper.

## INFORMATION FOR CANDIDATES

The approximate allocation of marks is given in brackets [ ] at the end of each question or part question.

You are advised that an answer may receive no marks unless you show sufficient detail of the working to indicate that a correct method is being used.

Final answers should be given to a degree of accuracy appropriate to the context.

The total number of marks for this paper is 60.

**This question paper consists of 3 printed pages and 1 blank page.**

1 The continuous random variable $X$, which takes all values in the interval $(-1, 1)$, has probability density function

$$f(x) = k(1 - x^2) \qquad -1 < x < 1$$

where $k$ is a constant.

(i) Show that $k = \frac{3}{4}$. Sketch the probability density function. [4]

(ii) Find the mean and variance of $X$. [5]

(iii) Find the probability that an observed value of $X$ will fall in the interval $(-0.1, 0.1)$. [2]

(iv) Find the probability that the mean of 100 independent observations of $X$ will fall in the interval $(-0.1, 0.1)$. [4]

2 The Reverend Thomas, a clergyman in the north of England who is also a keen statistician, has been monitoring the lengths of his sermons. He aims for each sermon to be between 10 and 15 minutes long, but in fact the sermons' lengths are given by the random variable $X$ which is Normally distributed with mean $13\frac{1}{2}$ minutes and standard deviation 2 minutes. The lengths of different sermons are independent of one another.

(i) Find the probability that an individual sermon lasts between 10 and 15 minutes. [3]

(ii) During a particular week, Rev. Thomas gives four sermons. Find the probability that their total length is more than an hour. [3]

(iii) Rev. Thomas is asked to provide a series of sermons to be broadcast in religious radio programmes but is instructed that he must reduce their length. Suppose he is successful to the extent that the random variable giving the sermons' lengths is now $\frac{1}{2}X$. Find the time interval required in a radio programme to ensure that, with probability 0.9, there is time for a sermon. [5]

(iv) Because of other variable elements in the radio programmes, the time available for a reduced-length sermon is itself a random variable, Normally distributed with mean 8 minutes and standard deviation 0.5 minutes. Find the probability that there is time for a sermon. [4]

3 A commuter's train journey to work is scheduled to take 52 minutes. Having noticed that he is *always* late, even when the trains are running normally, he decided to keep records for a random sample of ten journeys. On two of these occasions, there were major signal failures leading to severe disruption and complete suspension of services. He therefore decided to eliminate these two occasions from his records. On the other eight occasions, his journey times in minutes were as follows.

65 61 62 60 59 62 61 57

(i) Carry out a two-sided 5% test of the hypothesis that his overall mean lateness is 10 minutes. State the required distributional assumption underlying your analysis. [7]

(ii) Provide a 99% confidence interval for the mean journey time. Hence comment on the railway company's policy of offering refunds for journeys that are more than 15 minutes late. [6]

(iii) Comment on the commuter's decision not to include the two occasions when there were major signal failures. [2]

4 Part of a large simulation study requires the provision of many simulated observations that can be taken as coming from a distribution with mean 2. There is a suspicion that this part of the simulation is not working properly.

(a) 200 of the simulated observations are recorded in the form of a frequency table as follows.

| Simulated observation | 0 | 1 | 2 | 3 | 4 | $\geqslant 5$ |
|---|---|---|---|---|---|---|
| Frequency | 20 | 40 | 52 | 49 | 27 | 12 |

Carry out a $\chi^2$ test, at the 5% level of significance, to examine whether it is reasonable to suppose that the simulated observations come from the Poisson distribution with mean 2. [7]

(b) A further random sample of 200 simulated observations is taken. On this occasion, the variance of the underlying distribution is assumed to be 2. It is now required to test the null hypothesis that $\mu$, the mean of the underlying distribution, is 2 against the alternative hypothesis that $\mu > 2$. The usual test based on the sample mean $\bar{x}$ is to be used, with a 5% level of significance.

(i) State the size of the Type I error for this test. [1]

(ii) Show that the test rejects the null hypothesis if $\bar{x} > 2.1645$. [2]

(iii) Calculate the probability of a Type II error if in fact $\mu = 2.3$. [5]

# Mark Scheme

| | | | |
|---|---|---|---|
| 1(i) | $f(x)=k(1-x^2), \quad -1<x<1$<br>$1=\int_{-1}^{1}k(1-x^2)dx=k\left[x-\frac{x^3}{3}\right]_{-1}^{1}$<br>$=k\left[1-\frac{1}{3}-\left(-1+\frac{1}{3}\right)\right]=k\times\frac{4}{3}$<br>$\therefore \underline{\underline{k=3/4}}$<br>Sketch | Condone sloppy notation (e.g. omission of $dx$) throughout.<br>M1 Correct integral <u>used</u>, including limits (or a subsequent attempt to use them). Accept a correct alternative method using the c.d.f.<br>A1 Correctly integrated.<br>M1 Equated to 1, and their $k$ found convincingly. Beware printed answer. Allow max M1A0M1 for incorrect $k$. Accept convincing verification.<br>G1 Reasonable parabola: symmetrical about $y$-axis, with labels at intercepts $x=-1, 1$. | 4 |
| (ii) | $E[X]=\frac{3}{4}\int_{-1}^{1}(x-x^3)dx=\frac{3}{4}\left[\frac{x^2}{2}-\frac{x^4}{4}\right]_{-1}^{1}$<br>$=\frac{3}{4}\left[\frac{1}{2}-\frac{1}{4}-\left(\frac{1}{2}-\frac{1}{4}\right)\right]=0$<br><br>$Var(X)=E[X^2]=\frac{3}{4}\int_{-1}^{1}(x^2-x^4)dx$<br>$=\frac{3}{4}\left[\frac{x^3}{3}-\frac{x^5}{5}\right]_{-1}^{1}=\frac{3}{4}\left[\frac{1}{3}-\frac{1}{5}-\left(\frac{-1}{3}+\frac{1}{5}\right)\right]$<br>$=\frac{3}{4}\times\frac{4}{15}=\frac{1}{5}$ | M1 Integral correct (including limits as above). Condone $k$ missing or altered. Condone omission of or errors with notation "$E[X]$".<br>A1 c.a.o. Allow use of incorrect $k$ only if it f.t. from (i); result should still be 0. Award both marks for (convincing) "0 by symmetry".<br><br>M1 Use of formula for $Var(X)$. Possibly achieved as $E[X^2]-(E[X])^2$.<br>M1 Integral for $E[X^2]$ correct (including limits as above). Independent of previous M. Condone $k$ missing or altered.<br>A1F Condone omission of, or errors with, notation "$E[X^2]$". f.t. from c's $E[X]$ and/or $k$ from (i) provided variance is positive. | 5 |
| (iii) | $P(-0.1<X<0.1)=\frac{3}{4}\int_{-0.1}^{0.1}(1-x^2)dx$<br>$=\frac{3}{4}\left[x-\frac{x^3}{3}\right]_{-0.1}^{0.1}$<br>$=\frac{3}{4}\left[0.1-\frac{0.1^3}{3}-\left(-0.1+\frac{0.1^3}{3}\right)\right]$<br>$=\frac{3}{4}\left[0.2-\frac{2\times0.001}{3}\right]=0.1495$ | M1 Correct integral (including limits as above). Accept use of c's $k$.<br><br>A1 c.a.o. (Strictly) | 2 |

| | | | |
|---|---|---|---|
| (iv) | $\overline{X}_{100} \sim$ approx $N\left(0, \frac{1/5}{100} = 0\cdot002\right)$ | M1 For knowing to use Normal distribution (possibly ).<br>M1 For $\mu$ and $\sigma^2/n$, obtained from (ii). Must involve $n$ correctly. Variance must be > 0. | |
| | $P(-0.1 < \overline{X}_{100} < 0.1)$<br>$\approx P\left(\frac{-0\cdot1}{\sqrt{0\cdot002}} (= -2\cdot236) < N(0,1) < +2\cdot236\right)$<br>$= 2 \times 0\cdot4873$ (or $0\cdot9873 - 0\cdot0127$)<br>$= 0\cdot9746$ | M1 Formulating the problem as a 2-sided inequality and standardising. Accept equivalent forms e.g. based on symmetry. Be lenient w.r.t. abuse of inequalities; diagrams may well provide helpful clarification.<br>A1 c.a.o. Accept anything between 0·974 and 0·975 inclusive (but not a.w.r.t.).<br>Accept alternative method using $T = X_1 + \ldots + X_{100} \sim N(0, 20)$ to find $P(-10 < T < 10)$. | 4 |
| | | | **15** |
| **2(i)** | $X \sim N(13\tfrac{1}{2}, 2^2)$<br>$P(10 < X < 15) = P\left(\frac{10 - 13\frac{1}{2}}{2} < Z < \frac{15 - 13\frac{1}{2}}{2}\right)$ | M1 Award ONCE here or elsewhere for standardising. Condone $\mu - \bar{x}$ in numerator. | |
| | $= P(-1\cdot75 < Z < 0\cdot75)$<br>$= 0\cdot7734 - 0\cdot0401$ | A1 Both $\Phi$ values, not necessarily subtracted. Accept unsimplified forms, e.g. 1 – 0·9599 for 0·0401. | |
| | $= 0\cdot7333$ | A1 c.a.o. Expect 3 d.p. or better. | 3 |
| (ii) | $T = X_1 + X_2 + X_3 + X_4$<br>$\sim N\left(4 \times 13\tfrac{1}{2} = 54,\ \sigma^2 = 4 + 4 + 4 + 4 = 16\right)$ | B1 Mean.<br>B1 Variance (or $\sigma = 4$, provided it's clear). | |
| | $P(T > 60) = P\left(Z > \frac{6}{4} = 1\cdot5\right)$<br>$= 1 - 0\cdot9332 = 0\cdot0668$ | A1F f.t. incorrect mean but not variance. | 3 |
| (iii) | $\tfrac{1}{2}X \sim N\left(\tfrac{1}{2} \times 13\tfrac{1}{2} = 6\cdot75,\ \sigma^2 = (\tfrac{1}{2})^2 \times 4 = 1\right)$<br>Require $t$ such that $0\cdot9 = P\left(\tfrac{1}{2}X < t\right)$<br>$= P\left(Z < \frac{t - 6\cdot75}{1}\right)$ | B1 Mean.<br>B1 Variance (or $\sigma = 1$, provided it's clear).<br>M1 Formulation of the requirement as a one-sided inequality. | |
| | But $0\cdot9 = P(Z < 1\cdot282)$<br>$\therefore t - 6\cdot75 = 1\cdot282$<br>$\therefore t = 8\cdot032$ | B1 1·28(2) c.a.o.<br>SC After M0 for a two-sided inequality allow B1 for + and – 1·645 seen.<br>A1F f.t. incorrect mean (but not variance) or c's 1.282. | 5 |
| (iv) | Time available $Y \sim N(8, 0\cdot5^2)$<br>We want $P\left(Y - \tfrac{1}{2}X > 0\right)$ | M1 Formulation of the requirement as $P(Y -$ c's $\tfrac{1}{2}X > 0)$.<br>$P(Y >$ ans(iii)) scores 0/4. | |
| | $= P\left(N\left(1\cdot25,\ \sigma^2 = \frac{1}{4} + 1 = \frac{5}{4}\right) > 0\right)$ | B1F Mean. f.t. for 8 – mean (iii).<br>B1F Variance (or $\sigma$, provided it's clear). f.t. for 0·25 + Var (iii). | |
| | $= P\left(Z > \frac{-1\cdot25}{\sqrt{\frac{5}{4}}} = -1\cdot118\right) = 0\cdot868(1)$ | A1 c.a.o. | 4 |
| | | | **15** |

| Q | Answer | Mark | Notes | Total |
|---|---|---|---|---|
| 3(i) | $\bar{x} = 60 \cdot 875$<br>$s_{n-1}^2 = 5 \cdot 5536, \quad s_{n-1} = 2 \cdot 3566$ | | Allow $s_n^2 = 4 \cdot 8594, \ s_n = 2 \cdot 2044$ but ONLY if correctly used in sequel. | |
| | Assumption of underlying Normality. | B1 | For the assumption (about population, not data; e.g. do not allow just "it's Normal"). | |
| | Test statistic is $\dfrac{60 \cdot 875 - 62}{\dfrac{2 \cdot 3566}{\sqrt{8}}}$ | M1 | Numerator might be given as 8·875 – 10. Allow M1A0 then f.t. for $\mu - \bar{x}$ in numerator. Allow $s_n/\sqrt{7}$ (see above). | |
| | $= -1 \cdot 35(024)$ | A1 | c.a.o. Correct answer ww scores 2/2. | |
| | Refer to $t_7$ | M1 | May be awarded even if test statistic is wrong. Must see evidence of intention to use $t$-distribution. But no f.t. if $\nu$ is wrong. | |
| | Double tail 5% point is 2·365 | A1 | No f.t. if wrong. May be +ve or –ve. | |
| | Not significant. | B1 | For comparison (p.i.) and simple conclusion (p.i.) consistent with c's $t$ and critical value. | |
| | Accept hypothesis that mean lateness = 10 minutes. (Or that mean journey time = 62) | B1 | Consistent contextual conclusion. | 7 |
| | | SC | $t_8$ and 2·306 or $t_7$ and 1·895 used can score max B1 for either form of conclusion seen. | |
| (ii) | | N.B. | ZERO OUT OF 4 if not same distribution as used for test. Same wrong distribution can score max M1B0M1A0. | |
| | 99% C.I. given by: | | BUT allow recovery to $t_7$ for possible 4/4. | |
| | $60 \cdot 875 \pm 3 \cdot 499 \times \dfrac{2 \cdot 3566}{\sqrt{8}} = 60 \cdot 875 \pm 2 \cdot 915(3)$ | M1 | for $\bar{x} \pm$ ... Allow c's $\bar{x}$ from (i) or $\bar{x}$ – 52. | |
| | | B1 | for 3·499 (from $t_7$). | |
| | | M1 | for $s/\sqrt{8}$. Allow c's $s_{n-1}$ from (i). Also allow $s_n/\sqrt{7}$ (see above). | |
| | $= (57 \cdot 960, \ 63 \cdot 790)$ | A1 | BOTH. c.a.o. Accept correct ww for 4/4. Must be an interval. Min 2 dp required. | |

| $s$ | $t$ | 60.875+/- | lower | upper | | |
|---|---|---|---|---|---|---|
| 2.3566 | 3.499 | 2.91531 | 57.9597 | 63.7903 | Full marks | |
| 2.2044 | 3.499 | 2.72703 | 58.148 | 63.602 | M1B1M0A0 | |
| 2.3566 | 3.355 | 2.79533 | 58.0797 | 63.6703 | M1B0M1A0 if $t_8$ in (i) | |
| 2.2044 | 3.355 | 2.6148 | 58.2602 | 63.4898 | M1B0M0A0 if $t_8$ in (i) | |

(3·355 is $t_8$(1%).)

| Q | Answer | Mark | Notes | Total |
|---|---|---|---|---|
| | 15 minutes late corresponds to journey time of 67 minutes ...<br>• this is (well) outside the interval | E1 | A statistical comment: e.g. to include an explicit reference to the interval. | |
| | • so the penalty will hardly ever be invoked (despite regular lateness). | E1 | A contextual comment. Alternatively allow 2, 1, 0 for candidate who shows appreciation of the relationship of outliers to the mean. | 6 |
| (iii) | Seems reasonable to exclude these two occasions as they will not reflect normal daily conditions ... | E1 | Either "exclude" or "include", together with a reason. | |
| | ... but, strictly speaking, the sample is no longer a random one. | E1 | Recognise need for the sample to be random. Allow 2, 1, 0 for comments which address the distributional assumptions. | 2 |
| | Full credit for answers as outlined above. The real point, of course, consists of subtle but important discussions as to exactly what the underlying population is (all the journeys, or just the normal ones?). FULL CREDIT for discussing this. | | | |
| | | | | **15** |

4(a)

| $x$ | 0 | 1 | 2 | 3 | 4 | 5+ |
|---|---|---|---|---|---|---|
| Probab'y | 0.1353 | 0.2707 | 0.2707 | 0.1804 | 0.0902 | 0.0527 |
| Exp freq | 27.06 | 54.14 | 54.14 | 36.08 | 18.04 | 10.54 |
| Obs freq | 20 | 40 | 52 | 49 | 27 | 12 |
| Contrib'n | 1.8420 | 3.6930 | 0.0846 | 4.6266 | 4.4502 | 0.2022 |

| | | | | |
|---|---|---|---|---|
| | "Model" is $X \sim$ Poisson(2). | B2 | Expected frequencies. −1 e.e., but 2 errors and sum = 200 counts as only 1 error. Allow max B1if $e$'s do not add up to 200 (e.g. last $e = 7{\cdot}218$ leading to $X^2 = 17{\cdot}86$, for $X = 5$). Deduct 1 if $e$'s rounded to integers. | |
| | $X^2 = 14{\cdot}89(86)$ | B1F | f.t. from incorrect $e$'s (including when cells are combined as a consequence) provided Poisson (2) used/attempted. f.t. from here if incorrect. | |
| | Refer to $\chi^2_5$. | M1 | Allow f.t. from c's table. No f.t. from here if this does NOT agree with c's table i.e. $\nu$ = no. cells used − 1. Accept anything which implies the use of $\chi^2_5$, including LH tail. | |
| | Critical value at 5% is 11·07. | A1F | f.t. correct $\nu$ from c's table. For $\nu = 6$ the critical value is 12·59. No f.t. from here if incorrect critical point used. | |
| | Significant. | B1 | For comparison (p.i.) and preliminary conclusion (p.i.) consistent with c's $X^2$ and critical value. Accept equivalents e.g. "Accept $H_1$". | |
| | Seems Poisson(2) cannot be assumed. | B1 | For a consistent contextual conclusion. Require e.g. "Poisson(2)" or "required Poisson", not just "Poisson". | 7 |
| | | SC | $\chi^2_6$ and 12·59 or $\chi^2_4$ and 9·488 used can score max B1 for either form of conclusion seen. | |
| (b) (i) | 5% | B1 | | 1 |
| (ii) | Reject $H_0$ if $\bar{x} > 2 + 1{\cdot}645 \times \frac{\sqrt{2}}{\sqrt{200}} = 2{\cdot}1645$ | M1 | For $\bar{x}$ + (some point from N(0, 1)) $\times \frac{\sqrt{2}}{\sqrt{200}}$. | |
| | | A1 | For 1·645 leading to 2·1645 given in the question. | 2 |
| (iii) | If $\mu = 2{\cdot}3$, $\bar{X} \sim N\left(2{\cdot}3, \frac{2}{200} = 0{\cdot}01\right)$ | M1 | Distribution of $\bar{X}$, given $H_0$. | |
| | $P(\text{Type II error}) = P(\text{accept } H_0 \mid \mu = 2{\cdot}3)$ | M1 | Meaning of a Type II error in this context. | |
| | $= P(\bar{X} < 2{\cdot}1645 \mid \bar{X} \sim N(2{\cdot}3, 0{\cdot}01))$ | M1 | From (ii). | |
| | $= P\left(N(0, 1) < \frac{2{\cdot}1645 - 2{\cdot}3}{0{\cdot}1} = -1{\cdot}355\right)$ | A1 | $z$ value. (c.a.o. or f.t. c's variance - see how it goes.) | |
| | $= 1 - 0{\cdot}9123 = 0{\cdot}0877$ | A1 | c.a.o. | 5 |
| | | | | **15** |

# Examiner’s Report

## Statistics 3 (2615)

### General Comments

For this component there were 75 candidates from a total of 16 centres. The overall standard of these candidates was generally pleasing: many were well-prepared for it. The candidates were usually comfortable with the parts of questions involving process and calculation, but were less successful with the parts requiring comment or interpretation. Their explanations were often woolly and imprecise, indicating that they had not thought deeply enough about the issues which they were invited to consider.

Questions 1 and 2 were found to be particularly accessible and proved to be quite high scoring for well-prepared candidates. In Question 3 the work seen was quite disorganised and careless. Question 4 often started well but faded out in part (b).

Invariably all four questions were attempted but there was evidence to suggest that many candidates were running out of time when they came to Question 4(b), which may account for the poorer performances here.

### Comments on Individual Questions

**Question 1 (Continuous random variables; no context.)**

In this question the overall presentation, including the notation of definite integrals, could have been better.

(i) This part was well done with limits being used correctly and the integral equated to 1. It was surprising and disappointing to see some candidates producing either an incorrect sketch of the p.d.f. (usually triangular) or no sketch at all.

(ii) Most candidates answered this part correctly. Occasionally the integration went wrong but this was fairly rare. It was interesting to note that a number of candidates who, having obtained the value of the mean, then commented that the symmetry of the sketch confirmed their result. It would have been all the more pleasing if these candidates had had the courage to use this symmetry in the first place.

(iii) Many candidates set up the correct integral for this probability and most saw it through to the correct answer.

(iv) Only the best candidates coped well with this part. Many were unsure about the mean and variance of the sample mean and this led to a considerable amount of muddled work. Most knew that they had to use the Normal distribution. Candidates were seen to switch arbitrarily from $\bar{X}$ to $\Sigma X$ (which could give the same result) without seeming to appreciate the difference.

(ii) Mean 0, variance 1/5; (iii) 0·1495; (iv) 0·9746

**Question 2 (Linear combinations of Normal variables; sermon broadcast times.)**

(i) This part was almost always answered successfully.

(ii) This part too was usually correct. When an error occurred it was usually in the calculation of the variance of the required distribution.

(iii) The mean was almost always correct but an incorrect variance of 2 was commonly seen. Some candidates were unable to interpret the requirement correctly: they looked for a two-sided interval using a standardised Normal value of ±1·645 rather than 1·282.

(iv) There were relatively few completely correct answers to this part. Weaker candidates made no attempt at all. Others did not realise that they needed to combine the distribution of the broadcast time available with that of the reduced sermon time. These candidates tried to find the probability that the time available would exceed their answer to part (iii). Candidates who progressed beyond these two stumbling blocks still experienced difficulties obtaining the correct parameters and interpreting the requirement.

(i) 0·7333; (ii) 0·0668; (iii) 8·032; (iv) 0·8681

**Question 3 (Hypothesis test and confidence interval for population mean from a small sample; train journey times.)**

(i) It was clear that candidates knew in broad terms what was expected here, but all too often their work was spoilt by a lack of care and attention to detail and this frequently resulted in loss of marks. On many occasions the test statistic was incorrect because the wrong variance had been used. There was some confusion of journey time with lateness, and it was not uncommon to see the test statistic calculated as $t = (60{\cdot}875 - 52)/\ldots$ Care was needed also when it came to reading the tables of the $t$ distribution: a critical value of 1·895 instead of 2·365 was often seen. The inappropriate use of the Normal distribution cropped up occasionally.

The required distributional assumption was often either missing or imprecisely stated.

(ii) Candidates were able to demonstrate that they knew about confidence intervals, but unfortunately errors similar to those in part (i) were commonplace. Some candidates found the confidence interval for the mean lateness rather than the mean journey time. It was not uncommon to see the use of a different distribution from the one used in part (i).

Comments on the company's policy were expected to take the following line of argument. "15 minutes late corresponds to a journey time of 67 minutes. This is well beyond the confidence interval and so the mean journey time is very unlikely to exceed it. Hence, on average, the company is unlikely to have to pay many refunds."

An alternative response was to consider a journey time of 67 minutes in relation to the supposed population of individual journey times, and in particular, to discuss the extent to which 67 might be considered to be an outlier.

Most candidates thought informally along the lines of the former, but in the process exposed their lack of understanding of confidence intervals. The alternative response was rarely seen.

(iii) The comments made about the decision to exclude two items from the sample tended to concentrate on the supposed effect on the sample mean and the confidence interval. Many discussed whether or not the sample would be representative. Hardly any candidates showed the depth of insight to consider what the underlying population might be or the effect on the required distributional assumptions of altering the sample.

(i) Test statistic −1·350, critical value 2·365; (ii) (57·96, 63·79)

**Question 4 (Chi-squared test for goodness of fit of a Poisson model; a simulation study.)**

(a) In this part of the question fully correct working for the test statistic $X^2$ was seen regularly. Most remembered to find the frequency for *5 or more* rather than just 5. A number of candidates misjudged the number of degrees of freedom as 6 – 2 or 5 – 1 and hence looked up the wrong critical value. Usually the final conclusion was stated carelessly as "the model is not Poisson" rather than "the model is not Poisson(2)".

(b) At this stage some candidates were showing signs of running out of time. Candidates from some centres were obviously better prepared than others for this topic.

(i) When it was attempted this part gave no trouble at all.

(ii) There was a little uncertainty about what to do here. Some candidates used the value 2·1645 to work back to 1·645, but then they were likely to forget to explain its significance.

(iii) There was even more uncertainty here. Mistakes occurred at every stage of setting up the calculation: showing an understanding of the meaning of a type II error in this context, identifying the correct distribution to use given the true value of the mean, and using the critical value found in part (ii).

(a) Expected freq's 27·06, 54·14, 54·14, 36·08, 18·04, 10·54; $X^2 = 14·899$; $\nu = 5$, critical value 11·07; (b)(i) 5%; (iii) 0·0877