Curriculum Council

EXAMINERS' REPORT ON 2001 TERTIARY ENTRANCE EXAMINATION

SUBJECT: APPLICABLE MATHEMATICS

STATISTICS

Year	Number Who Sat	Non-Examination Candidates	Did Not Sit
2001	4577	38	182
2000	4783	72	204
1999	4787	49	221

The Examiners' Report is written by the Chief Examiner (or another examiner on their behalf) to comment on matters relating to the Tertiary Entrance Examination in their subject. The opinions and recommendations expressed in this report are those of the Chief Examiner and not necessarily representative of or endorsed by the Curriculum Council.

The Marking Guide provided at the end of this report was prepared for markers and substantially amplified by discussions held in the pre-marking meeting. It is not intended as a set of model answers, and is not exhaustive as regards alternative answers. Some of the answers are less than perfect, but represent a standard of response that the examiners deemed sufficient to earn full marks. Teachers who use this guide should do so with its original purpose in mind.

SUMMARY/ABSTRACT

The paper, covering the majority of the syllabus, had a mean score of 64.5% (again above the range of 55 - 60% suggested by the Curriculum Council) reflecting a slightly easier paper than the 2000 and 1999 papers (means of 63.5% and 60.58% respectively). It should be noted that the examiners had tried to set a slightly harder paper than the previous two years. Whilst the paper resulted in better discrimination in the top range of marks, it failed in lowering the overall average mark. The fourth use of graphics calculators and two A4 sheets of notes again gave no perceived problems and the paper seemed generally to be well accepted. A good range of marks was achieved with markers once again reporting very few scripts with really low marks and, in general, competency by the candidates. As in previous years, the combination of routine (part) questions with more challenging ones discriminated across the range of candidate abilities, especially for the top candidates. The write-on Question/Answer booklet format continued to work extremely well giving the candidates no apparent problems, but significantly aiding the marking process.

GENERAL COMMENTS

This paper, the fourth examination of the syllabus allowing the use of graphics calculators and candidates to take sheets of notes into the examination, was handled well by the candidates. The majority of the candidates' papers indicated that there had been sufficient time to attempt all questions and hardly any had working on the "spare" pages. The examiners tried to arrange the questions in approximate order of ease and not worth, making it difficult to judge whether, when questions were not attempted, this was due to lack of time or knowledge.

As stated in the summary, the overall average percentage mark for the paper was higher this year. However, there were four questions with average marks below 50% (lowest 34.1%) whilst the lowest average percentage mark for a question was 52.92% in 2000. The lowest number of candidates attempting a question was 3675 compared to 3665 in 2000, despite 206 fewer candidates in 2001.

Once again, recurring comments in the markers' reports relate to the candidates lack of ability to **read** the questions carefully ('thousands' missed in Q4 and Q10, time not given in Q3 and Q11, no comments in Q17) and subsequently not answering what is actually asked. Candidates are still not giving coherent **explanations** and/or **justifications** when asked for them. Lack of units and inappropriate numbers of **decimal places** were again very common. Candidates continue to have problems with **time periods** in regression models (Q10 and Q15). This year initial values were given, but candidates obviously have problems with the arbitrariness of this, indicating it was impossible to have negative values. (There is no reason why the years 1991 to 1999 could not be represented by time periods -9 to -1 by using $t = 0$ in 2000. However, this would generally not be done!) Also converting a calculated time period back to a specific year caused problems for many candidates.

The candidates' answers again revealed competency with the graphics calculators, but there are still some problems both in interpreting results and entering data. Just quoting the calculators error message when two matrices cannot be multiplied (Q1) or not checking for errors in data entry for regression questions (Q10 and Q15). Also many candidates failed to graph over the complete domain (12 months) in Q9. There was again little evidence this year of any differences between the use of the various calculator models. However, it should be noted that, once again, the examiners tried very hard to eliminate possible differences through choice of wording and values.

Despite the extra stress placed on markers due to the short marking time-line, the marking was achieved speedily and cheerfully. The continued use of the optical mark reader sheets and the sorting of the sheets into individual folders at the Curriculum Council was, once again, very much appreciated. The examiners wish to thank all those, markers and Curriculum Council staff, who took part in the marking process.

COMMENTS ON SPECIFIC QUESTIONS

1 worth 8 marks attempted by 4575 candidates mean mark = 7.34 = 91.8%

This question, attempted by all, but two candidates, was very well done and had the highest mean percentage mark. Apart from the very occasional numerical error, those candidates that lost marks did so because they failed to indicate that the matrix multiplication in (ii) could not be done due to the number of columns in the first matrix D not equalling the number of rows in the second matrix G. Far too often the candidates made some general statement about wrong dimensions or reported the error given by their calculators. This type of lack of precision in explanations has been noted in previous years, both in matrix multiplication and addition/subtraction.

2 worth 5 marks attempted by 4538 candidates mean mark = 4.32 = 86.3%

This question had the second highest average mark and was very well done. The main error occurred in part (a) with candidates either giving the total number of cards (80), instead of the number in each deck, or using the column totals instead of the row totals and getting varying numbers in each of the four decks.

3 worth 5 marks attempted by 4487 candidates mean mark = 3.31 = 66.2%

This question on the normal distribution was both 11th in popularity and mean mark. The markers were appalled by the large number of candidates who failed to give the answer to part (a)(i) $P(X = 15)$ as zero. Two common answers were 0.5 and 0.1974. The first suggests a misuse of the calculator by just using 15 and obtaining $P(X \leq 15)$. The second suggests a misunderstanding of the correction for continuity by trying to use it on an already continuous distribution and calculating $P(14.5 \leq X \leq 15.5)$. Many candidates failed to indicate the time in part (b).

4 worth 9 marks attempted by 4492 candidates mean mark = 7.25 = 80.5%

Whilst being 10th in popularity, this question was 5th in average mark and generally well done. The common mistakes were in missing out the thousands in (a) and using the wrong denominator (number of visitors instead of nights spent by visitors) in (b). Whilst the following did not arise out of mistakes made by candidates but subsequent discussion about the paper, it should be noted that for any subject covered by the Mathematics (TES) Syllabus Committee that the meaning of the word "**average**" (as used in "calculate the average number") is always the "**arithmetic mean**".

5 worth 10 marks attempted by 4534 candidates mean mark = 7.01 = 70.1%

On the whole, this question was well done. Only a few candidates used the mean instead of the median for the box and whisker plot. (It should be noted that the median box plot is the only one on the syllabus even though calculators give other ones.) A common error in (b) was to use 16.5 – 6 instead of 16.5 – 6/40. The common error in (c) was the failure to realise how the sums of squares would change or even that it would change.

6 worth 6 marks attempted by 4568 candidates mean mark = 4.96 = 82.6%

This question was very well done being 3rd in popularity and 4th in mean value. Two common mistakes were giving the required probability in (a)(ii) as 0 and failing to justify with numerical values why events *A* and *B* are not independent.

7 worth 9 marks attempted by 4557 candidates mean mark = 6.95 = 77.3%

Another question that was well done with very few errors in the tree diagram in (a) (just a few candidates gave only those specifically listed eg only Q following P and missing R). However, the main mistake was in (c) in interpreting the request for the "combination (in any order) of CDs" as the single selection R followed by Q.

8 worth 6 marks attempted by 4518 candidates mean mark = 4.97 = 82.9%

Although 8th in popularity, this question was very well done, being 3rd in average mark. Just a handful of candidates gave a general word description rather than a description in the form requested.

9 worth 5 marks attempted by 4115 candidates mean mark = 3.01 = 60.1%

Surprisingly badly done, especially as this type of question, requiring the solution of inequalities via the calculator, has been pretty standard for the previous three years. It was 21st (out of 22) in popularity and 17th in mean mark. Those candidates who did attempt the question often found only the first or second (3rd was required) points of intersection of the two functions in the required domain. Many candidates had difficulty translating the point (correct or incorrect) in time to an approximate date.

10 worth 12 marks attempted by 4537 candidates mean mark = 8.59 = 71.6%

This question was moderately done. After problems on the last two years' papers, the examiners decided to give the candidates the starting point for the time variable. Despite this, some candidates still used other starting points, but failed to indicate what these were. As the data is given to 2 decimal places, the coefficients in (a) really needed to be given to at least three decimal places. Many candidates rounded which caused inaccuracies later on in the question. In (c), many candidates again forgot the thousands (Q4 also) and erroneously indicated that prediction at one time point outside the data was extrapolation and therefore not at all accurate. Whilst prediction several time points outside the data is not at all reliable, prediction one time point away is reasonable especially when, as shown in (b), the correlation coefficient is high. In part (d), many candidates indicated that prediction was inappropriate simply because the time value would be negative and therefore impossible – failing to see the arbitrariness of the starting point for time in problems like this. In part (e), many candidates failed to comment upon the "number" of visitors, commenting instead on what the graph looked like or what was happening to the residuals. 'Thousands' were often missing from part (g).

11 worth 11 marks attempted by 4438 candidates mean mark = 6.82 = 62.0%

Some candidates used areas of triangles whilst others used integrals. Either method was appropriate. Several candidates had problems finding *b* in (a) and many forgot to give the time (Q3 also) in (d).

12 worth 9 marks attempted by 4572 candidates mean mark = 5.68 = 63.1%

Whilst 2nd in popularity this question had the 13th average mark. Candidates had little problem with (a) and (b), but little idea about (c). Many candidates found the correct denominator for the probability, but generally did not think of all possibilities for the numerator with $\binom{6}{1}\binom{4}{1}\binom{16}{1}$ being the commonly given erroneous value.

13 worth 9 marks attempted by 4507 candidates mean mark = 5.72 = 63.5%

Parts (a) and (b) were generally well done. In (c), candidates tended to give matrices that either preserved the proportion or doubled the area but did not do both. In (d), most candidates used a specific matrix to show the required situation rather than trying for general values. Many who did try to use general values ended up giving confusing arguments.

14 worth 13 marks attempted by 4463 candidates mean mark = 7.95 = 61.1%

Many candidates handled the exponential probability distribution in (c) and (d) better than the Poisson distribution in (a) and (b). The large variety of means that were used for both the Poisson and exponential distributions were interesting! Justification for any number of ships (correct or incorrect) in part (b) was extremely poor.

15 worth 10 marks attempted by 4470 candidates mean mark = 6.13 = 61.3%

Again the examiners gave the starting point for time and most candidates used this. However, despite being asked for the "exponential" relationship, several candidates gave a linear relationship – not always the log-linear equivalent model. Too few decimal places were also a problem. Most candidates were unable to work out the average percentage increase in (b) just using 4.15 from the power of *e*!! In (c), candidates again erroneously decided that one time period was extrapolation and would yield an unreliable prediction. Unfortunately, some candidates who thought that the prediction would be acceptable did so because "the correlation coefficient was high". **It is very important to note that the correlation coefficient is a measure of LINEARITY and NOT appropriate for an EXPONENTIAL model.** (Whilst the calculators do give a value for *r* with the exponential calculations this should be treated with extreme caution. On some calculators, it is the appropriate value for the log-linear equivalent model, but the HP 38G only gives *r* for a linear fit even if it does perform an exponential fit.) In (d), many candidates had problems working out that the $t = 10.9$ value corresponded to the 11th census that would take place in the year 2016.

16 worth 8 marks attempted by 4443 candidates mean mark = 3.76 = 46.9%

Candidates often made errors in row reduction calculations, some even attempting illegal operations such as adding the same value to each term in a row. Unfortunately, even those candidates who achieved the correct reduced form were unable to show clearly that there was always at least one solution. Many only tried to deal with the "one" solution case rather that show that "no solution" was not possible. However, even those who failed to obtain a correct matrix in (a) generally managed to find a conditions on $\boldsymbol{p}$ for one solution.

17 worth 4 marks attempted by 4399 candidates mean mark = 2.13 = 53.25%

This question was not at all well done with a strong suggestion that candidates were not reading the question. Many candidates just found the numbers in the appropriate sets for (a) and (b) (which was NOT asked for) and completely failed to comment on the deliveries to the houses whose numbers were contained in the set.

18 worth 10 marks attempted by 4380 candidates mean mark = 6.74 = 67.4%

Candidates generally handled the binomial distribution in (a) and normal distribution in (b) well, but often failed to "indicate clearly" which distribution they were using. A few candidates had problems with the correction for continuity in (b). Unfortunately, many candidates were unable to complete (c). Those candidates who found the appropriate mean (some tried to use the mean from (b)) often then tried to use the normal approximation to the Poisson rather than the straight Poisson probability.

19 worth 14 marks attempted by 4396 candidates mean mark = 10.05 = 71.8%

Whilst 17th in popularity, this linear programming question was generally well done and had the 7th highest mean mark. Parts (a) and (b) were extremely well done. Most candidates handled (c) well but, unfortunately, in (d), some candidates forgot that short trains could make twice as many journeys as long trains when deciding, on potential income grounds, whether to use short or long trains.

20 worth 5 marks attempted by 4339 candidates mean mark = 1.81 = 36.2%

This question was quite different to standard time series questions, requiring interpretation rather than calculations, and was very badly done. The candidates' answers suggested almost a rote application of methods rather than consideration of the bigger picture. Most candidates indicated that the centred moving average line was increasing with time in (a), but very few candidates recognised the non-linear nature of the line by commenting on the change around time periods 13 to 17. (Do candidates realise to look for linearity in the CMAs?) In (b), some candidates thought a single model would be appropriate for the time periods 3 to 23 because of random residuals (from where??!!). Those that thought a single model was not appropriate invariably gave no valid reason. A few brave candidates drew a straight line in on the diagram and then saw that the CMAs would be above for the first half and below for the second half and hence a single model was not at all appropriate. (I would have loved to give these candidates bonus marks for truly exploring the situation.) A few other candidates suggested a linear followed by an exponential relationship.

21 worth 4 marks attempted by 3675 candidates mean mark = 1.36 = 34.1%

This question was lowest in popularity and had the lowest average mark. This question was designed to test candidates' ability with matrix algebra. However, it was also useful to remember that in the given context of systems of linear equations A^{-1} only exists for a unique solution, which is clearly not the scenario here. Hardly any candidates showed any coherent working in (a) with most missing the hint of "a" (as opposed to "the") solution (let alone the mention of other solutions in (b)) and trying to use inverse matrices. By sheer luck, most candidates correctly guessed one of the many answers to (b). Fortunately for them, no justification was asked for as hardly any candidates gave any.

22 worth 8 marks attempted by 4356 candidates mean mark = 3.04 = 38.1%

Whilst most candidates correctly worked out in (a) the total guests' bill and the total for uneaten food, far too many candidates failed to express this as a proportion. (Failure to read again?!) Many candidates managed to get three or four correct equations in (b) (sometimes in their own undefined variables!!), but failed miserably in (c). (Although there were more than four equations possible in (b), it did help to have the appropriate combination in (c).) Some enterprising candidates tried to pull the wool over the marker's eyes in (c) by conjuring an equation out of thin air that quickly reduced to $n = 1$!

POINTS FOR CONSIDERATION BY THE SYLLABUS COMMITTEE

It may well be appropriate to formally insert into the syllabus the use of the Poisson distribution as an approximation for the binomial distribution when n is large and p is very low (giving a small mean). For Applicable Mathematics candidates, this is possibly a more useful approximation than the normal approximation to the Poisson.

An extra stress on the use of the correlation coefficient ONLY with linear regression would not go astray.

Possibly not a syllabus matter, but more a teaching matter: once again the examiners' stress the importance of READING questions carefully and answering the questions that are asked.

The format of the paper, use of calculators and sheets of notes seem to be working well and should be maintained.

Jen Bradley
December 2001

2001 Examining Panel
Chief Examiner: Mrs Jennifer Bradley
Chief Marker: Mr Laurie Sutton
Third Member: Dr David Wilson

Chief Marker: Mr Laurie Sutton

1. (8 marks)

For the matrices

$$D = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}, \quad E = \begin{bmatrix} 1 & -2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}, \quad F = \begin{bmatrix} 1 & -3 \\ 2 & 6 \end{bmatrix} \text{ and } G = \begin{bmatrix} 0 & -4 \\ -1 & 5 \end{bmatrix}$$

calculate the following, giving detailed reasons if the calculation is not possible.

(i) $2F - G$ [2]

$\begin{bmatrix} 2 & -2 \\ 5 & 7 \end{bmatrix}$	-1 1st error -1/2 subs round up	**2**

(ii) DG [2]

not possible number cols in D ≠ number rows in G	1 1	**2**

(iii) ED [2]

$\begin{bmatrix} 8 \\ 32 \\ 24 \end{bmatrix}$	-1 1st error -1/2 subs round up	**2**

(iv) F^{-1} [2]

$\frac{1}{6-(-6)}\begin{bmatrix} 6 & 3 \\ -2 & 1 \end{bmatrix} = \frac{1}{12}\begin{bmatrix} 6 & 3 \\ -2 & 1 \end{bmatrix} \left\{= \begin{bmatrix} 1/2 & 1/4 \\ -1/6 & 1/12 \end{bmatrix}\right\}$ $\left\{= \begin{bmatrix} 0.5 & 0.25 \\ -0.167 & 0.083 \end{bmatrix}\right\}$	-1 1st error -1/2 subs round up	**2**

TOTAL 8

2. (5 marks)

Ash, Brock, Chris and Di have each put together a deck of cards containing types Earth, Fire, Air and Water. In defining the rules of the game, they have decided that each Earth card is worth ***e*** points, each Fire card is worth ***f*** points, each Air card is worth ***a*** points and each Water card is worth ***w*** points. The distribution of cards and the total number of points in each of the four decks are represented by the following matrix system of equations.

$$\begin{bmatrix} 5 & 5 & 5 & 5 \\ 5 & 7 & 4 & 4 \\ 5 & 6 & 5 & 4 \\ 4 & 4 & 4 & 8 \end{bmatrix} \begin{bmatrix} e \\ f \\ a \\ w \end{bmatrix} = \begin{bmatrix} 900 \\ 910 \\ 910 \\ 840 \end{bmatrix}$$

(a) How many cards are there in each deck? [1]

20	1	**1**

(b) Write down the inverse of the coefficient matrix given above. [2]

$\begin{bmatrix} 2.8 & 1 & -3 & -0.75 \\ -1.2 & 0 & 1 & 0.25 \\ -1.2 & -1 & 2 & 0.25 \\ -0.2 & 0 & 0 & 0.25 \end{bmatrix}$	-1 1st error -1/2 subs round up	**2**

(c) Find the point values for each type of card. [2]

$\begin{bmatrix} e \\ f \\ a \\ w \end{bmatrix} = \begin{bmatrix} 70 \\ 40 \\ 40 \\ 30 \end{bmatrix}$ earth worth 70 points fire worth 40 points air worth 40 points water worth 30 points	-1 1st error -1/2 subs round up **but** $\begin{bmatrix} 70 \\ 40 \\ 40 \\ 30 \end{bmatrix}$ only 1	**2**

TOTAL 5

3. (5 marks)

A teacher travels to school each day by car. She sets out from home each day at 8 am exactly. The time taken for her journey can be considered to be normally distributed with a mean of 15 minutes and a standard deviation of 2 minutes.

(a) Calculate the probability that her journey takes

(i) exactly 15 minutes, [1]

Let X be time taken for journey $X \sim N(15, 4)$ $P(X = 15) = 0$	 1	 **1**

(ii) more than 18 minutes. [2]

$P(X > 18)$ $\{= P(Z > 1.5)\}$ $= 0.06681$	1 1	 **2**

(b) Find the time such that there is a 90% chance that the teacher will arrive before this time.

[2]

Let time after 8 am be T mins $P(X < T) = 0.9$ $\{P(Z < (T - 15)/2) = P(Z < 1.2815)\}$ $T = 17.563$ Time is 17 mins and 34 secs past 8 am (8.18 OK)	 1 1	 **2**

TOTAL 5

4. (9 marks)

The following data from the Australian Bureau of Statistics gives the number of nights, in thousands, spent by international visitors to Australia in 1999 by State (first six entries in table) or Territory.

NSW	Vic	Qld	SA	WA	Tas	NT	ACT
38334	20494	24928	4542	12349	1694	3044	2161

(a) If there were 4.459510 million international visitors to Australia in 1999, calculate the average number of nights spent in Australia per international visitor. [3]

$\frac{107546000}{4459510}$ $= 24.1$	1 1 1	**3**

(b) For a randomly chosen international visitor night in 1999, calculate the probability that it was

(i) spent in WA, [2]

$\frac{12349}{107546} \quad \{= 0.1148\}$	1 1	**2**

(ii) spent in one of the Territories (NT or ACT), [2]

$\frac{3044 + 2161}{107546} \quad \left\{= \frac{5205}{107546} = 0.0484\right\}$	1 1	**2**

(iii) not spent in NSW nor Qld. [2]

$1 - \frac{38334 + 24928}{107546} \quad \left\{= 1 - \frac{63262}{107546} = 0.4118\right\}$	1 , 1	**2**

TOTAL 9

5. (10 marks)

A data set, X, consists of 40 data items, x_i. Some summary values for the data set are given below.

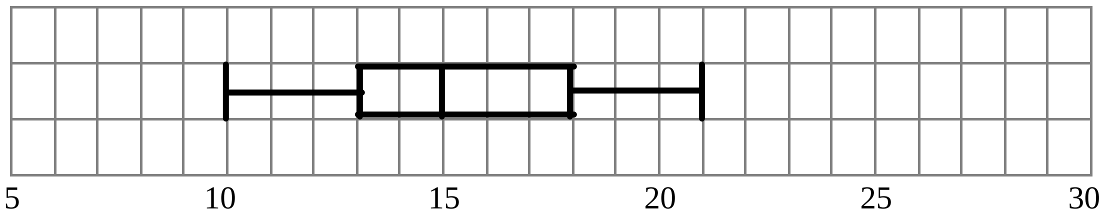
mean	median	standard deviation
16.5	15	2.5

minimum	maximum	1st quartile	3rd quartile
10	27	13	18

Unfortunately, when the data was recorded the largest data item was written down as 27 instead of 21.

For the **correct** data values

(a) draw an appropriate box and whisker plot on the grid below, [3]



See above	-1 each error	**3**

(b) write down the value of the mean, [3]

$16.5 - \dfrac{27-21}{40}$	1 , 1	
$= 16.5 - 6/40 = 16.35$	1	**3**

(c) give the value of the variance, if the original $\sum_{i=1}^{40} x_i^2 = 11140$. [4]

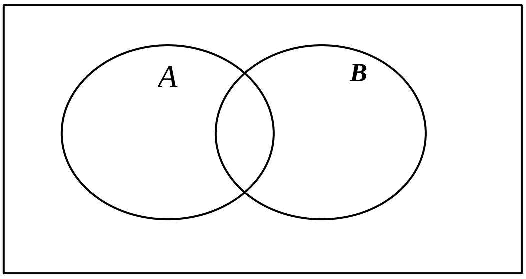
New $\sum x_i^2 = 11140 - 27^2 + 21^2 = 10852$	2	
Variance = $\dfrac{\sum x_i^2}{n} - \bar{x}^2 = \dfrac{10852}{40} - 16.35^2$	1	
$= 3.9775$	1	**4**

TOTAL 10

6. (6 marks)

For the events A and B represented in the Venn diagram below

$\mathrm{P}(A \cap B) = 0.3$, $\mathrm{P}(A) = 0.6$ and $\mathrm{P}(A|B) = 0.75$.



(a) Find

(i) P(B), [2]

$\mathrm{P}(A\mid B)=\dfrac{\mathrm{P}(A\cap B)}{\mathrm{P}(B)}$ $\quad\therefore\quad$ $\mathrm{P}(B)=\dfrac{\mathrm{P}(A\cap B)}{\mathrm{P}(A\mid B)}$	1	
$\mathrm{P}(B)=\dfrac{0.3}{0.75}=0.4$	1	**2**

(ii) $\mathrm{P}\overline{(A \cup B)}$. [2]

$\mathrm{P}\overline{(A\cup B)} = 1 - \mathrm{P}(A\cup B) = 1 - [\ \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A\cap B)]$	1	
$= 1 - 0.6 - 0.4 + 0.3$ $= 0.3$	1	**2**

(b) Are events A and B independent? Justify your answer. [2]

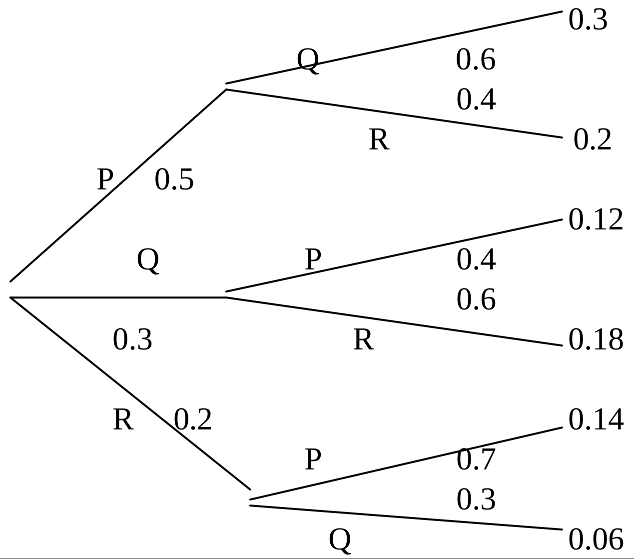
$\mathrm{P}(A)\mathrm{P}(B) = 0.6\times 0.4 = 0.24 \neq 0.3 = \mathrm{P}(A\cap B)$ **OR** $\mathrm{P}(A) = 0.6 \neq 0.75 = \mathrm{P}(A\mid B)$	1	
Events A & B are NOT independent	1	**2**

TOTAL 6

7. (9 marks)

Kim has three favourite CDs: P, Q and R. He always plays two of them whilst completing his homework each night. There is a 50% chance that he will put CD P on first each evening and a 20% chance that he will put CD R on first. At any time there is a 60% chance that he will follow CD P by CD Q, a 40% chance that he will follow CD Q by CD P and a 70% chance that he will follow CD R by CD P. Kim never plays the same CD twice.

(a) Display the above information in a tree diagram, indicating clearly the probabilities for playing each CD and the probabilities for each possible order for playing the CDs. [4]



1 first branch	
2 second branch	
1 final probs	

(b) What is the probability that Kim will not play CD Q ? [2]

{P & R + R & P} = 0.2 + 0.14 {= 0.34}	1 , 1	**2**

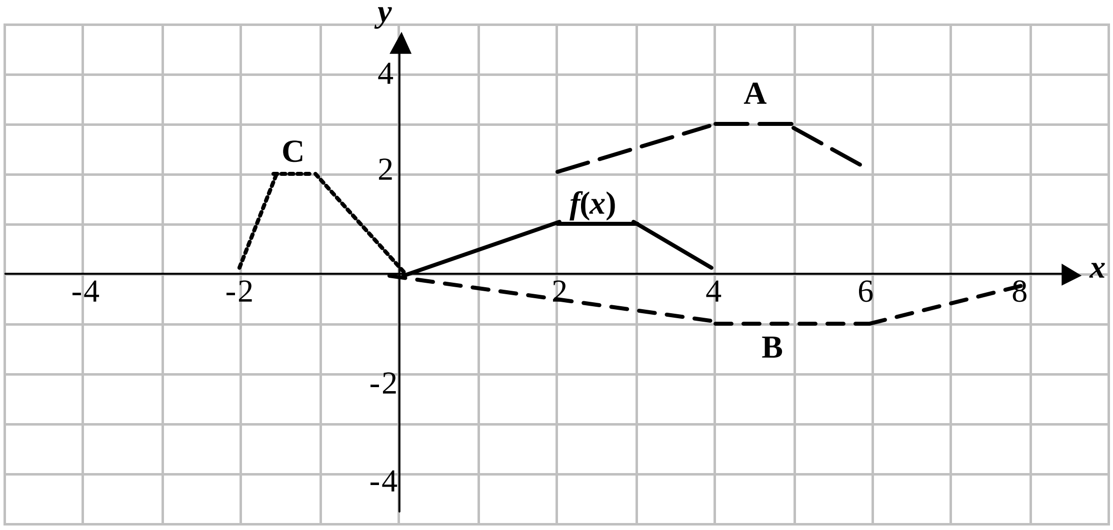
(c) What is the least probable combination (in any order) of CDs? Justify your answer. [3]

P & Q 0.3 + 0.12 = 0.42	1	
P & R 0.34		
Q & R 0.18 + 0.06 = 0.24	1	
Least probable combination is Q & R	1	**3**

TOTAL 9

8. (6 marks)

The graph of $y = f(x)$ is shown as a solid line on the graph below. Describe each of the graphs A, B and C (shown as dotted lines) as transformations of $y = f(x)$ in the form $y = \boldsymbol{a}\, f(\boldsymbol{b}x + \boldsymbol{c}) + \boldsymbol{d}$.



A: $y = f(x - 2) + 2$	1 , 1	
B: $y = -f(x/2)$	1 , 1	
C: $y = 2\, f(-2x)$	2 but -1 for any of the 3 parts wrong (no -ve mark)	**6**

TOTAL 6

9. (5 marks)

In the wake of a series of policy disasters and unfavourable economic results, the government of a small country is finding its popularity is in decline and the popularity of the opposition appears to be increasing. Polling conducted for the government suggests that its popularity index over the next few months will follow the function

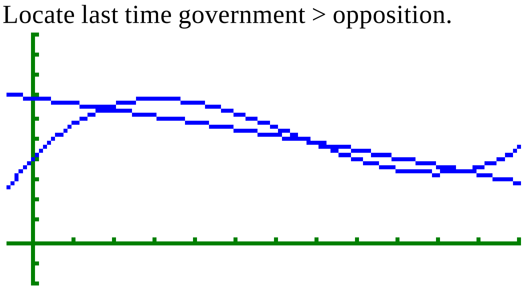
$$\ell n\left(2-\frac{t}{18}\right), \qquad \text{where } t \text{ is the number of months since June } 1^{st}\ 2001.$$

During the same time period the opposition's popularity index is thought to be slightly more volatile following the function

$$\frac{2}{5}+\frac{149}{720}t-\frac{7}{160}t^2+\frac{29}{12960}t^3 .$$

The Prime Minister must call an election before June 2002 but, naturally, he wants to remain in office for as long as possible. What is the latest date that he can set for the election so that the government will still hold a popularity advantage over the opposition?

Explain clearly how you arrived at your answer.

Draw the two functions on GC & automatically find points of intersection. Locate last time government > opposition.  { $-0.6 \le t \le 12$ $\quad -0.2 \le y \le 1.0$ }	2	
Intersections t = 1.78, 6.893 & 10.619	1	
Last point is appropriate. Election before 10.619 months.	1	
Call election for on or before April 18^{th} 2002 { "Mid April" or anywhere between April 16 to April 19 OK}	1	
		5

TOTAL 5

10. (12 marks)

The table below contains data given by the Australian Bureau of Statistics (ABS) relating to the number of international visitors to Australia for the years 1991 to 1999.

Year	Number of International Visitors (in thousands)
1991	2370.40
1992	2603.27
1993	2996.22
1994	3361.72
1995	3725.83
1996	4164.83
1997	4317.87
1998	4167.21
1999	4459.51

(a) Calculate the least squares regression line for predicting the number, n, of international visitors in time period t, where $t = 0$ in 1990.

[2]

$n(t) = 274.911166t + 2199.53972$ { ≥ 3 dp ≥ 3 dp}	1 , 1	**2**

(b) Calculate the correlation coefficient, r, between the number of international visitors and time.

[1]

$r = 0.9688$	1	**1**

(c) Predict the number of international visitors in 2000 and comment upon the accuracy of your prediction.

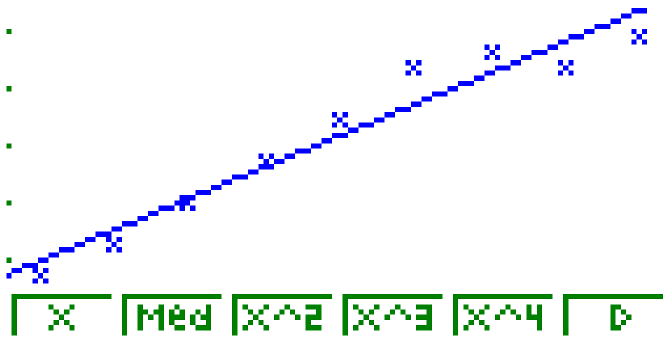
[2]

$t = 10$ number = 4,948,651 reasonably accurate {r high & only one year away}	1 1	**2**

(d) Explain why it would or would not be appropriate to use your regression line to predict how many international visitors there were in 1985.

[2]

NOT appropriate 6 years outside range of data - extrapolation	1 1	**2**



A plot of your regression line on a scatter diagram of the data should look similar to that given above.

(e) Comment upon the number of international visitors in 1998 and 1999 (the last two years) compared to previous years.

[2]

There were not as many visitors as expected.	2	**2**

Using the data for 1991 to 1997 an appropriate regression line is

$n = 346.255t + 1977.85714$ with n and t defined as previously.

(f) How many fewer international visitors were there in 1998 than would have been predicted by the 1991 to 1997 model?

[2]

$346.255 \times 8 + 1977.85714 = 4747.89714$ $4747.89714 - 4167.21 = 580.68714$ 580,687 fewer visitors than predicted	1 1	**2**

(g) The correlation coefficient between the number of international visitors and time for the 1991 to 1997 model is 0.9961. If the data had been expressed in millions (eg 2.37040 million in 1991) and the time period had been defined using t = 1990 in the year 1990, what would be the new correlation coefficient?

[1]

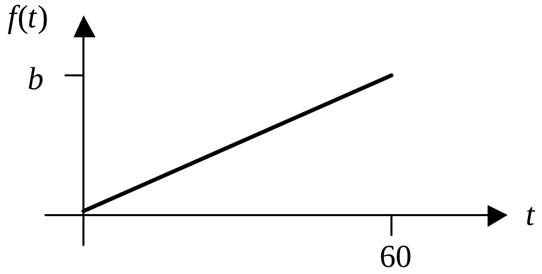
0.9961 {same value OK}	1	**1**

TOTAL 12

11. (11 marks)

Lee never arrives at school before 8.00 am and never arrives after 9.00 am. The probability distribution function for her time, t, of arrival at school is given below, with constant b and t being measured as minutes after 8.00 am.

$$f(t) = \frac{b}{60}t$$



(a) What is the value of b ? [2]

$\frac{1}{2} \times 60 \times b = 1$ $\quad b = 1/30$	1 , 1	**2**

(b) Calculate the probability that Lee arrives at school

(i) before 8.30 am, [2]

$\frac{1}{2} \times 30 \times 30/1800 = \frac{1}{4} = 0.25$	1 , 1	**2**

(ii) after 8.40 am. [2]

$1 - \frac{1}{2} \times 40 \times 40/1800 = 1 - 4/9 = 5/9 = 0.55556$	1 , 1	**2**

(c) Given that Lee arrives at school after 8.30 am, what is the probability that she arrives after 8.40 am?

[2]

$\frac{\text{P(time} > 8.40)}{\text{P(time} > 8.30)} = \frac{5/9}{3/4} \left\{ = \frac{20}{27} = 0.7407 \right\}$	1 1	**2**

(d) Find the time, T , to the nearest minute, such that the probability that Lee arrives at school before T am is the same as the probability that she arrives at school after T am.

[3]

P(time > t) = P(time < t) = 0.5 where T = 8. t am $\frac{1}{2} \times t \times t/1800 = \frac{1}{2}$ $\quad t^2 = 1800 \quad t = 42.43$ time is 8.42 am to nearest minute	1 1 1	**3**

TOTAL 11

12. (9 marks)

(a) Complete the table below to show the number of student councillors at City Community College.

[2]

-1 1st error -1/2 subs round up

	Male	Female	Total
Year 12	6	**4**	10
Year 11	**3**	5	8
Total	**9**	**9**	**18**

(b) Calculate the probability that a randomly selected student councillor

(i) is a female, [1]

9/18 {= 0.5}	1	**1**

(ii) is in Year 12 given that the councillor is male. [2]

$\frac{6}{9}$ {= 0.6667}	1 1	**2**

(c) Three student councillors are selected at random. Find the probability that this selection includes at least one male from Year 12 and at least one female from Year 12.

[4]

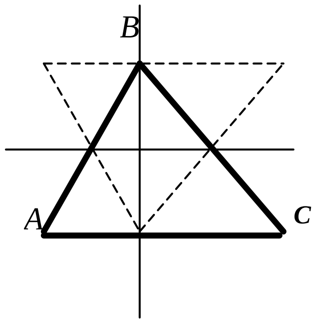
$\dfrac{\text{M12 F12 Y11 + M12 F12 F12 + M12 F12 M12}}{\text{TOTAL}}$ $= \dfrac{\binom{6}{1}\binom{4}{1}\binom{8}{1} + \binom{6}{1}\binom{4}{2} + \binom{6}{2}\binom{4}{1}}{\binom{18}{3}}$ $\left\{\dfrac{6\times4\times8+6\times6+15\times4}{816} = \dfrac{288}{816} = 0.3529\right\}$	1 , 1 , 1 1	**4**

TOTAL 9

13. (9 marks)

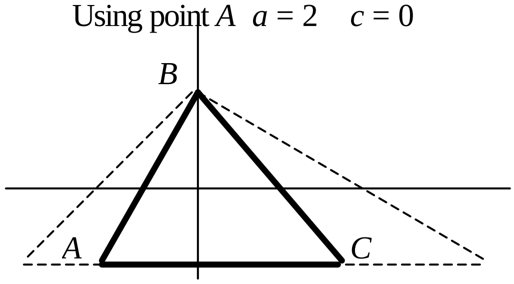
The Bouncy Rubber Company (BRC) are setting up a corporate Website into which they want to incorporate an animated graphic using their triangular logo. The BRC logo can be modelled in a cartesian plane as the triangle with vertices A at $(-2, -2)$, B at $(0, 2)$ and C at $(3, -2)$.

(a) For the first part of the animation the BRC want the logo to turn upside down. Write down a transformation matrix that will map the standard logo to an upside down one. [2]



Reflect about x axis $\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$ accept rotation 180° $\begin{bmatrix}-1 & 0\\0 & -1\end{bmatrix}$ $\left\{\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}\begin{bmatrix}-2 & 0 & 3\\-2 & 2 & -2\end{bmatrix}=\begin{bmatrix}-2 & 0 & 3\\2 & -2 & 2\end{bmatrix}\right\}$	2 or nothing	**2**

(b) For the second part of the animation the BRC want to stretch the logo horizontally so that the original triangle is transformed to a triangle with vertices $(-4, -2)$, $(0, 2)$ and $(6, -2)$. Write down an appropriate transformation matrix. [2]

$\begin{bmatrix}a & b\\c & d\end{bmatrix}\begin{bmatrix}-2 & 0 & 3\\-2 & 2 & -2\end{bmatrix}=\begin{bmatrix}-4 & 0 & 6\\-2 & 2 & -2\end{bmatrix}$ Using point B $b = 0$ $d = 1$ Using point A $a = 2$ $c = 0$  $\begin{bmatrix}2 & 0\\0 & 1\end{bmatrix}$	-1 1st error -1/2 subs round up	**2**

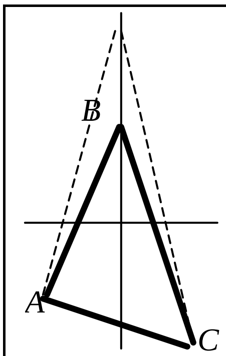
(c) At another stage in the animation the BRC want the logo to appear double in area (while remaining in proportion). Write down a transformation matrix which will double the area.

[2]

New area triangle = 2× (½ base × height) In proportion ½ × $\sqrt{2}$ base × $\sqrt{2}$ height $\begin{bmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix}$	1 mark preserving proportion 1 mark doubling area	**2**

(d) Suppose the vertex *C* was originally at (2, −3) instead of (3, −2). Show clearly that it is not possible to use a matrix transformation to achieve a stretch in the vertical direction which keeps the vertices *A* and *C* fixed but moves the vertex *B* to (0, 4).

[3]



$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} -2 & 0 & 2 \\ -2 & 2 & -3 \end{bmatrix}=\begin{bmatrix} -2 & 0 & 2 \\ -2 & 4 & -3 \end{bmatrix}$ Using *B* $b = 0$ $d = 2$ then using *A* $a = 1$ $c = -1$ but using *C* $a = 1$ $c = 3/2$ Values of *c* are inconsistent so it is not possible to use a matrix transformation. accept $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} -2 & 0 & 2 \\ -2 & 2 & -3 \end{bmatrix}=\begin{bmatrix} -2 & 0 & 2 \\ \mathbf{-4} & 4 & \mathbf{-6} \end{bmatrix}$ which does not keep *A* & *C* fixed.	 1 1 1	**3**

TOTAL 9

14. (13 marks)

A small port has berthing for two container ships. As it generally takes about two days to unload each container ship the port handles on average 5 container ships a week with at most 7 container ships per week.

It is reasonable to think of the number of ships arriving in any given week as following a Poisson distribution. It is also reasonable to think of the time between one container ship and the next arriving as following an exponential distribution with a mean of 1.4 days.

(a) Calculate the probability that in any given week

(i) exactly 3 container ships arrive, [1]

Let X be number container ships arriving. $X \sim$ Poisson $m = 5$ per week $P(X = 3) = 0.14037$	1	**1**

(ii) at most 4 container ships arrive, [2]

$P(X \leq 4) = 0.44049$	1 , 1	**2**

(iii) more container ships arrive than can be handled. [3]

$P(X > 7) = 1 - P(X \leq 7)$ $= 1 - 0.86662 = 0.13338$	1 , 1 1	**3**

(b) What is the maximum number of container ships per week that the port should be able to handle so that there is less than a 2% chance that more container ships arrive than can be handled? Justify your answer.

[3]

$P(X > 8) = 1 - P(X \leq 8) = 0.0681$ $P(X > 9) = 1 - P(X \leq 9) = 0.03183$ $P(X > 10) = 1 - P(X \leq 10) = 0.0137$	1	
The last value is less than 2% = 0.02	1	
Port should be able 10 ships per week	1	

		3

Suppose that the port is empty and then one container ship arrives.

(c) Calculate the probability that the next container ship will arrive within the next day.

[2]

Let Y be time to next ship arriving Y ~ exponential $m = 1.4$ days $\{k = 1/1.4 = 5/7\}$ $P(Y < 1) = 1 - e^{-5/7}$ $\{= 0.51046\}$	1 , 1	**2**

(d) Calculate the probability that it will be more than two days before the next container ship arrives.

[2]

$P(Y > 2) = e^{-10/7}$ $\{= 0.23965\}$	1 , 1	**2**

TOTAL 13

15. (10 marks)

The table below gives the median age (in years) of the Australian population in the given years (with time period t) in which a census was carried out.

Census Year	t	Median Age
1966	1	26.6
1971	2	27.6
1976	3	28.7
1981	4	29.9
1986	5	31.3
1991	6	32.7
1996	7	34.0

(a) Calculate the equation for the exponential regression curve which best fits the above data.

[2]

Age = $25.4196184\, e^{0.04150669\, t}$ …1 $= 25.4196184 \times 1.042380135^{t}$ …2	1 , 1	**2**

(b) What is the average percentage increase in the median age for each five year period?

[2]

$\{a\, e^{bt} : 100(e^{b} - 1);\quad a\, b^{t} : 100(b - 1)\}$ 4.2%	2	**2**

(c) Predict the median age for 2001 and comment upon the accuracy of your prediction.

[2]

$t = 8$ age = 35.4 Reasonably accurate {only one time period away}	1 1	**2**

(d) It is estimated that when the median age reaches 40 years, Australia's working population will be too small to support those individuals on a pension. If the present trend continues, in what census year will the median age first exceed 40 years? Indicate clearly how you arrived at your answer.

[4]

From equation 1 $t = \ell n\,(40 / 25.4196184) / 0.04150669 = 10.9$ From equ 2 $t = \ell n\,(40 / 25.4196184) / \ell n\,(1.042380135) = 10.9$ OR draw graph 11th census so year 2016	Working 3 1	**4**

TOTAL 10

16. (8 marks)

Given the following system of equations for the variables x, y and z.

$$\begin{array}{ccccccccc} 2x & - & 3y & + & z & = & 4 \\ x & - & y & + & \boldsymbol{p}z & = & 3 \\ x & - & 2y & + & 2z & = & \boldsymbol{p}^2 \end{array}$$

(a) Show clearly that the system of equations always has at least one solution.

[6]

$\left[\begin{array}{ccc\|c} 2 & -3 & 1 & 4 \\ 1 & -1 & \boldsymbol{p} & 3 \\ 1 & -2 & 2 & \boldsymbol{p}^2 \end{array}\right]$ For no solution need 0 0 0 #	1	
$R_3 + R_2 \Rightarrow \left[\begin{array}{ccc\|c} 2 & -3 & 1 & 4 \\ 1 & -1 & \boldsymbol{p} & 3 \\ 2 & -3 & \boldsymbol{p}+2 & \boldsymbol{p}^2+3 \end{array}\right]$	1	
$R_3 - R_1 \Rightarrow \left[\begin{array}{ccc\|c} 2 & -3 & 1 & 4 \\ 1 & -1 & \boldsymbol{p} & 3 \\ 0 & -0 & \boldsymbol{p}+1 & \boldsymbol{p}^2-1 \end{array}\right]$ {other working possible}	1	
If $1 + \boldsymbol{p} = 0$ then $\boldsymbol{p} = -1$	1	
But then $\boldsymbol{p}^2 - 1 = 0$ => 0 0 0 0 => infinitely many solutions	1	
So "no solution" is not possible => at least one solution	1	
		6

(b) Indicate clearly the conditions required on $\boldsymbol{p}$ so that the system of equations has only one solution.

[2]

Need 0 0 @ #	1	
$1 + \boldsymbol{p} \neq 0 \quad \therefore \quad \boldsymbol{p} \neq -1$	1	
		2

TOTAL 8

17. (4 marks)

A street contains 20 houses numbered from 1 to 20. Three students deliver advertising leaflets to the houses. To make their delivery a little interesting they decide to deliver to the houses whose numbers are contained in one of the three sets given below. Huia delivers according to set *H*, Jan delivers according to set *J* and Leong delivers according to set *L*.

$H = \{ 2, 4, 6, 8, 10, 12, 14, 16, 18, 20 \}$
$J = \{ 3, 5, 7, 11, 13, 17, 19 \}$
$L = \{ 3, 5, 6, 9, 10, 12, 15, 18, 20\}$

(a) Comment on the deliveries at the houses whose numbers are contained in the set

$$(H \cap J) \cup (H \cap L) \cup (J \cap L) \ .$$

[2]

$\{ 6, 10, 12, 18, 20\} \cup \{3, 5\}$ Numbers in more than one set Receive two lots of advertising leaflets	 1 1	 **2**

(b) If *T* is the set containing all the numbers from 1 to 20, comment on the deliveries at the houses whose numbers are contained in the set

$$T \cap \overline{(H \cup J \cup L)} \ .$$

[2]

$\{ 1, 2, \ldots, 20\} \cap \{1\} = \{1\}$ Numbers not in any set Receive no advertising leaflets	 1 1	 **2**

TOTAL 4

18. (10 marks)

Joe Striker plays soccer with the Western Wonders soccer team. When his side wins a corner kick, Joe's teammates try to kick the ball to him 40% of the time.

(a) If the Western Wonders win 8 corner kicks in a game, calculate the probability that Joe's teammates try to kick the ball to him exactly 3 times. Indicate clearly how you perform your calculations.

[3]

Let X (Y & W) be number of times ball kicked to Joe $X \sim$ binomial $n = 8$ $p = 0.4$ $P(X = 3) = 0.27869$	 1 , 1 1	 **3**

(b) Over a period of one year, the Western Wonders won 96 corner kicks when Joe was playing. Use a normal approximation to find the probability that his teammates tried to kick the ball to Joe at least 40 times. Indicate clearly which distribution you are using.

[4]

$X \sim$ binomial $n = 96$ $p = 0.4$ $Y \sim$ normal $m = 38.4$ $s = 4.8$ $P(X \geq 40) \approx P(Y > 39.5) = 0.40936$ { $P(Z > 0.23) = 0.4090$ from tables} {NOTE: exact binomial = 0.40688 & $P(Y > 40) = 0.36944$}	1 , 1 1 , , 1 {-2 , -1}	 **4**

(c) When Joe tries to "head" a goal he is not very successful and Club records show that in fact he only has a 0.08 chance of success. With the probability of success so low, the Poisson distribution can be used to approximate the number of successful shots. (The mean is the same as that for a normal approximation.) Use the Poisson approximation to find the probability that, in his next 100 attempts to head a goal, Joe will be successful more than 7 times.

[3]

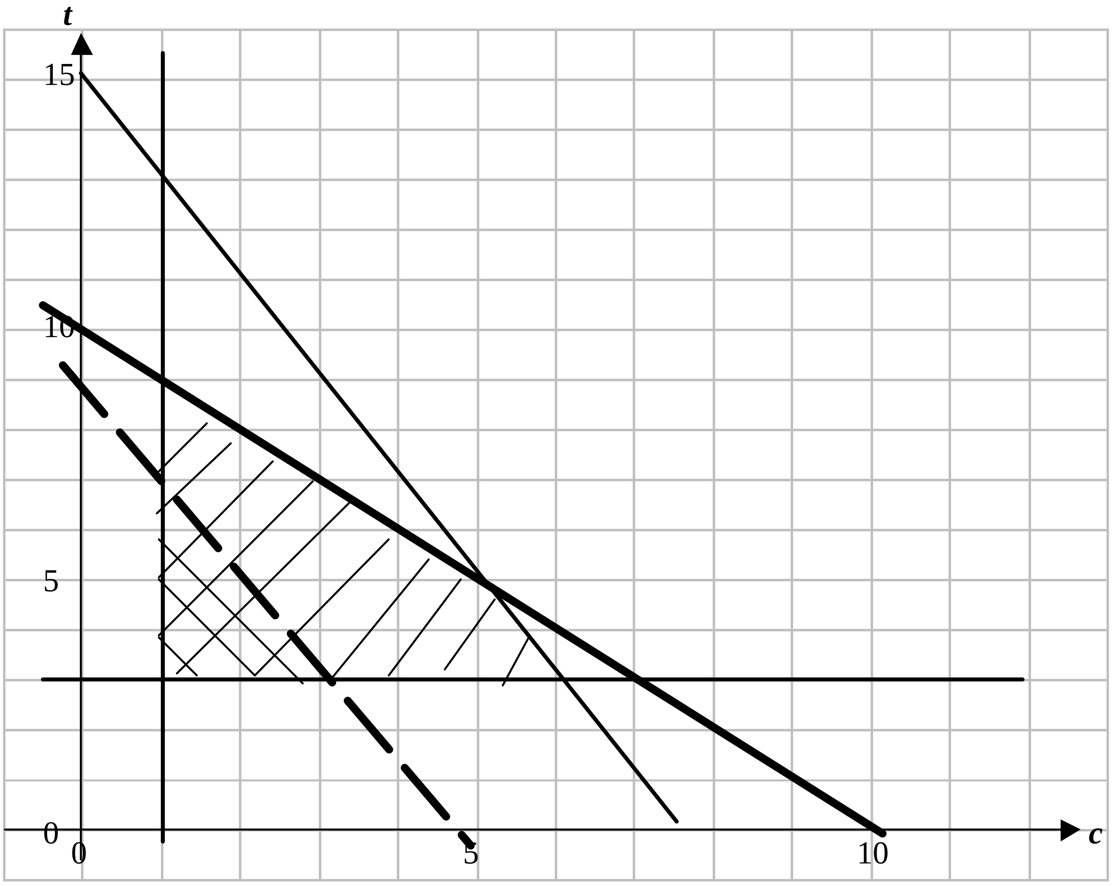
$X \sim$ binomial $n = 100$ $p = 0.08$ $W \sim$ Poisson $m = 8$ $P(W > 7) = 1 - P(W \leq 7)$ $= 0.54704$ {NOTE: exact binomial =0.5529}	 1 1 1 {-2}	 **3**

TOTAL 10

19. (14 marks)

The controller of a small steam railway has been examining the mix of passenger and goods services on each of the daily trains. To meet minimum service provisions each train must have at least one passenger coach and at least three goods trucks. To comply with safety standards each engine must not pull a total of more than 15 truck-lengths where one goods truck is one truck-length and one passenger coach is two truck-lengths. Each engine requires 0.5 tonne of coal to make each journey by itself plus 0.1 tonne of coal for each passenger coach or goods truck that it pulls but each engine can only carry 1.5 tonnes of coal.

Let c be the number of passenger coaches and t be the number of goods trucks per train. Some of the constraints on c and t have been drawn on the diagram below.



(a) Draw the remaining constraint(s) on the above diagram and shade in the feasible region. [3]

Constraints already drawn: $c > 1$ $t > 3$ $2c + t \leq 15$ Need $0.1c + 0.1t \leq 1$ Draw line shade region	 1 1 , 1	 **3**

(b) If each passenger coach has the potential to earn the railway $100 per journey and each goods truck has the potential to earn the railway $65 per journey, what is the maximum potential income per journey that can be earnt by the railway? What mix of passenger coaches and goods trucks is required to achieve this maximum? Justify your answer.

[5]

potential income = $100c + 65t$	1	
Point PI	2 working	
(1, 3) 295		
(1, 9) 685 {gradient method OK}	-1 each point missing except (1, 3)	
(5, 5) 825		
(6, 3) 795		
Maximum potential income of $825	1	
with 5 each coaches & trucks	1	
		5

The controller realises that if the trains were shorter the engine could travel faster and make twice as many journeys at no extra cost.

(c) If the train is now limited to a maximum of 9 truck-lengths, indicate clearly on the diagram the new feasible region and give the best mix of carriages in terms of potential income.

[4]

New constraint $2c + t \le 9$	1	
Draw line & shade region (cross hatched area)	1	
Point PI		
(1, 3) 295	1 working	
(1, 7) 555		
(3, 3) 495		
Best mix 1 passenger coach & 7 goods trucks	1	
		4

(d) Explain briefly, in terms of potential income, whether the railway should use short or long trains.

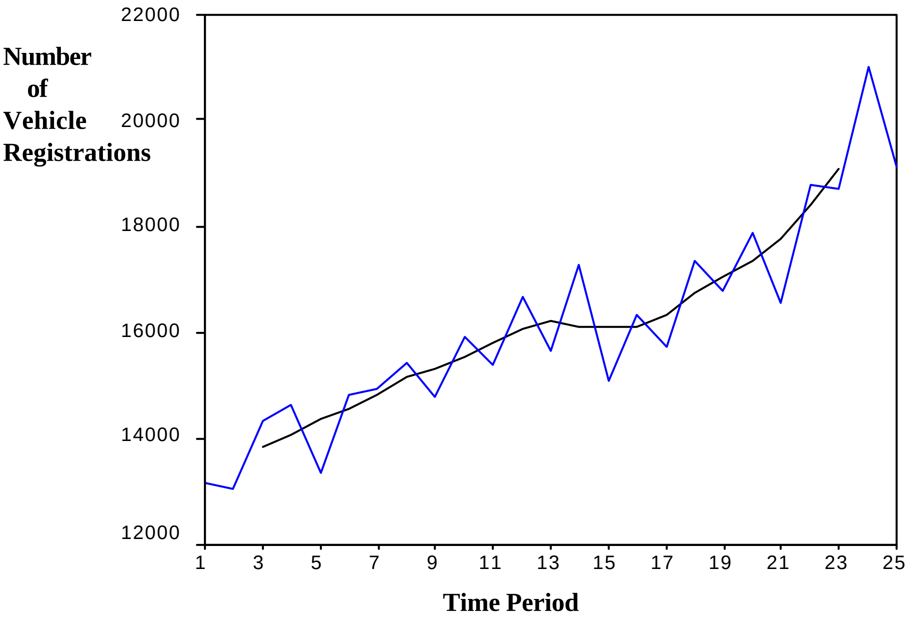
[2]

$2 \times 555 > 1 \times 825$	1	
Railway should use short trains	1	
		2

TOTAL 14

20. (5 marks)

The time series graph below indicates the number of vehicle registrations in Western Australia (zig-zag line) together with moving averages (smooth line) for each quarter from March 1992 to March 1998 (time periods 1 to 25). (The graph was produced from raw data from the Australian Bureau of Statistics). As the data is quarterly data a four-point centred moving average has been used.



(a) Describe any patterns shown by the centred moving average line. [2]

On the whole numbers are increasing But there is a dip in the middle {linear – dip – linear} {or linear - exponential or linear - quadratic}	1 1	**2**

The usual full model for prediction is given by

model = trend + seasonal component ,

where **trend** is the least squares regression line obtained from the centred moving averages.

(b) Comment upon the appropriateness of using this model for all or part of the time period from March 1992 to March 1998.

[3]

ONE trend line for whole NOT appropriate One line to about Dec 94 then another from Dec 95 to March 98 would be OK	1 1 1	**3**

TOTAL 5

21. (4 marks)

A system of three linear equations in three unknowns is represented by the matrix equation

$$AX = B, \qquad \text{where } X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \text{ and the matrix } B \text{ is \textbf{not} zero.}$$

(a) If $X_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ is a solution to the equation, explain clearly why $2X_1 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$ cannot also be a solution.

[2]

$AX_1 = B \quad \therefore \quad 2AX_1 = 2B = A2X_1$ IF $2X_1$ is a solution then $A2X_1 = B$ so $2B = B$	1	
Thus $B = 0$ BUT B is NOT zero so $2X_1$ is NOT a solution.	1	
		2

(b) If $X_2 = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$ is a second solution to the equation, find a third solution.

[2]

$AX_1 = B$ and $AX_2 = B$ $AX_1 + AX_2 = A(X_1 + X_2) = 2B$	1	
$\therefore \quad A\dfrac{X_1 + X_2}{2} = B$ so $\dfrac{X_1 + X_2}{2} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$ is a solution.	1	
$\left\{ \text{also } \begin{bmatrix} k \\ k+1 \\ k+2 \end{bmatrix} \text{ for any } k \right\}$		**2**

TOTAL 4

22. (8 marks)

Bruce and Kylie invited 186 guests to their wedding reception, which had a three- course meal (entree, main course and dessert).

37	guests did not eat an entree
180	guests ate a main course
164	guests ate a dessert
136	guests ate all three courses

(a) An entree cost \$6, a main course cost \$11 and a dessert cost \$5. If Bruce and Kylie were charged \$22 for each of their guests, what proportion of the bill was for uneaten food?

[2]

$1-\dfrac{149\times6+180\times11+164\times5}{186\times22}$ OR $\dfrac{37\times6+6\times11+22\times5}{186\times22}$ { 0.0973}	1 1	**2**

Let
eo be the number of guests who ate only the entree,
mo be the number of guests who ate only the main course,
do be the number of guests who ate only the dessert,
em be the number of guests who ate only the entree and the main course,
ed be the number of guests who ate only the entree and the dessert,
md be the number of guests who ate only the main course and the dessert,
n be the number of guests who ate nothing.

Thirteen guests ate only one of the courses, ie $eo + mo + do = 13$.

(b) Write down **four** more equations in the above variables, explaining briefly how you obtained them. (Hint: A Venn diagram might be useful.)

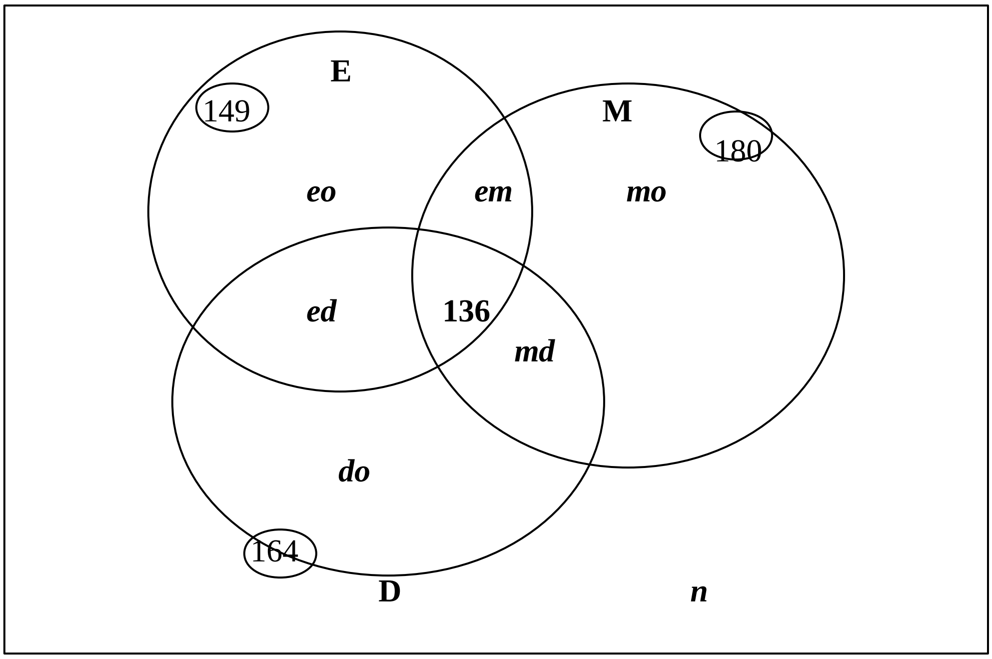
[4]

Entrée $eo + em + ed + 136 = 149$ or $eo + em + ed = 13$ Main $mo + em + md + 136 = 180$ or $mo + em + md = 44$ Dessert $do + ed + md +136 = 164$ or $do +ed + md = 28$ All $n + eo + mo + do + em + md + ed +136 = 186$ or $n + 13 + em + md + ed = 50$ or $n + em + md + ed = 37$	1 1 1 1	**4**

(c) Deduce that one guest ate nothing. [2]

Adding 1st 3 above eo + mo + do + 2em +2ed + 2md = 85	1	

$\therefore$ $em + ed + md = (85 - 13)/2 = 36$ & $n + 36 = 37$ so $n = 1$	1	**2**



Other equations for (b):

$mo + md + do + n = 37$

$em + ed + md + n = 37$

$eo + ed + do + n = 6$

$eo + em + mo + n = 22$