X202/13/01

NATIONAL QUALIFICATIONS 2012

FRIDAY, 25 MAY
1.00 PM – 4.00 PM

APPLIED MATHEMATICS
ADVANCED HIGHER
Statistics

Read carefully

1. Calculators may be used in this paper.
2. Candidates should answer all questions.

   Section A assesses the Units Statistics 1 and 2
   Section B assesses the Unit Mathematics for Applied Mathematics
3. **Full credit will be given only where the solution contains appropriate working.**
4. A booklet of Statistical Formulae and Tables is supplied for all candidates.

SQA ©

Section A (Statistics 1 and 2)

Marks

Answer all the questions

A1. There is a 20% chance of rain at the site where the *S in the Meadow* one-day festival is held annually on the first Saturday in September. A local weather station is able to predict rain correctly on 90% of days when it actually rains and incorrectly predicts rain on 5% of days when it does not rain. Given that the weather station forecasts rain, calculate the probability that it actually does rain at the festival. **4**

A2. In August 2010, on the first anniversary of the release of the man convicted of the 1988 Lockerbie bombing, a telephone poll of 1013 members of the general public, carried out for Scottish Television, revealed that 539 disagreed with the Scottish Justice Secretary's decision to release Abdelbaset Ali al-Megrahi and allow him to return to Libya on compassionate grounds.

(*a*) State the type of sampling used and indicate a possible disadvantage. **2**

(*b*) Stating any assumption required, calculate an approximate 95% confidence interval for the proportion of the general public that disagreed with the decision. Comment on the result, in relation to the statement on the website of the organisation carrying out the poll that over half of Scots think that the Scottish Justice Secretary was wrong. **5**

A3. A helicopter with a payload of 900 kg is used to transport groups of 8 climbers into a wilderness area. The weights (kg) of climbers are N(80,16) and the weights (kg) of their equipment packs are N(30,4). Stating any assumptions required, calculate the probability that the payload is exceeded. **6**

A4. It is known from experience that a machine used to fill soft drinks bottles of various sizes delivers "fill volumes" that are normally distributed with standard deviation 0·1 ml. In order to meet statutory requirements for fill volumes, it is desirable that the mean fill volume for 500 ml size bottles should be 500·30 ml. Shortly after the start of a production run of 500 ml bottles, a random sample of 10 bottles was taken and the fill volumes were:

500·40 500·22 500·31 500·13 500·16 500·18 500·32 500·31 500·18 500·44

(*a*) Use a statistical test to assess whether there is any evidence that the mean fill volume differs from 500·30 ml. **5**

(*b*) Under what circumstance would a *t*-test of the same hypothesis be required and how might this affect the conclusion? **2**

Marks

A5. In order to test, at the 5% level of significance, the null hypothesis, $\rho = 0$, versus the alternative, $\rho \neq 0$, when a sample of 25 pairs of (X, Y) values are available, write down the critical region in the form $|t| > t_c$ and hence show that it is equivalent to $|r| > 0{\cdot}40$. **4**

A biology student investigated the yield (Y) from crops of radishes from 25 seed trays with varying planting density (X). He obtained a correlation of –0·07 and concluded that there was no evidence of a relationship between X and Y. State what you would advise him to do before concluding that there is no relationship. **2**

[Turn over

Marks

A6. In 1982 David Raup and John Sepkoski published data in an article entitled *Mass Extinctions in the Fossil Record* in the journal *Science*. The table lists the number of recorded extinctions of marine invertebrate families in 76 intervals of time, all of equal duration.

Extinctions	0	1	2	3	4	5	6	7	8	9
Frequency	0	13	15	16	7	10	4	2	1	2

Extinctions	10	11	12	13	14	15	16	17	18	19	20
Frequency	1	1	0	0	1	0	2	0	0	0	1

The scientists compared the frequency distribution of extinctions to that expected from a Poisson distribution using a chi-squared goodness-of-fit test.

(*a*) State the null and alternative hypotheses. **1**

In carrying out the test, a statistician completed the table of observed and expected frequencies shown.

Extinctions	*Observed*	*Expected*
≤ 1	13	5·88
2	15	10·00
3	16	14·03
4	7	14·77
5	10	12·44
6	4	8·73
7	2	5·25
8	1	2·76
≥ 9	8	2·15

(*b*) Given that the mean number of extinctions per period is 4·21 verify that the expected frequency of one or fewer extinctions in a period is 5·88, showing all steps in your working, and state why two sets of frequencies have been amalgamated. **3**

(*c*) Given that the chi-squared statistic is 37·57, complete the test and state what may be concluded concerning the randomness of the extinctions of marine invertebrate families. **3**

Marks

A7. In a project carried out by pupils at Kingswood School in Bath supervised by Head of Mathematics, Mr Garrod Musto, and archaeologist, Mr Leslie Cram, measurements were made of a dog pawprint on a Romano-British tile in a local museum. The length of the pawprint examined was 71 mm but shrinkage occurs during tile firing so, in order to determine the size of the dog that made the print, the students mimicked Roman production methods for tiles and arranged for a sample of modern breeds of dog to make specimen prints on tiles prior to firing. The following data was obtained and a scatter plot suggested that a linear fit was appropriate.

Breed	Paw length (mm) from fired tile, x	Weight (kg), y
G. Shep. cross	69	27·40
Greyhound	61	26·45
Lurcher cross	43	11·65
Rott. cross	66	17·40
Labrador	65	30·30
St. Bernard	93	43·80
G. Shepherd	68	33·15
Staff. cross	48	16·20
Labrador	64	33·55

In the usual notation: $\overline{x} = 64{\cdot}111$, $\overline{y} = 26{\cdot}656$, $S_{xx} = 1592{\cdot}9$, $S_{xy} = 995{\cdot}04$ and $S_{yy} = 817{\cdot}67$.

(*a*) Show that the slope of the least squares regression line of y on x differs significantly from zero at the 1% level of significance. **6**

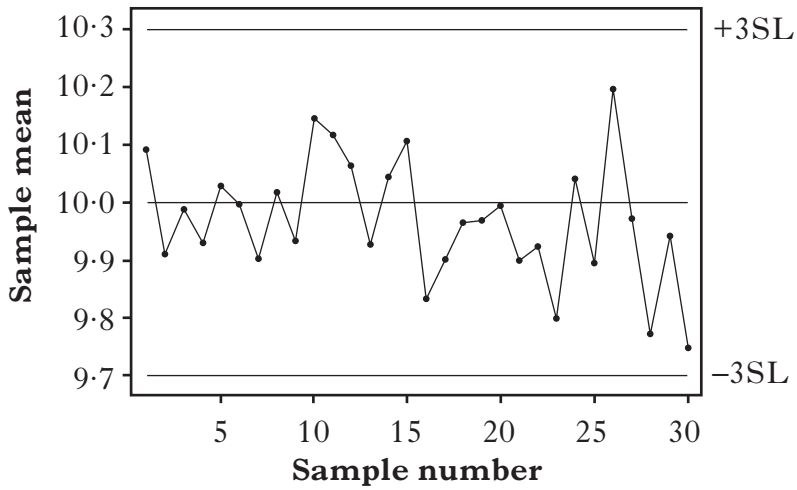
(*b*) Use the regression line equation to predict the weight of the dog that made the pawprint on the tile from the museum. **2**

(*c*) State any other tasks you might perform in analysing the given data. **2**

[Turn over

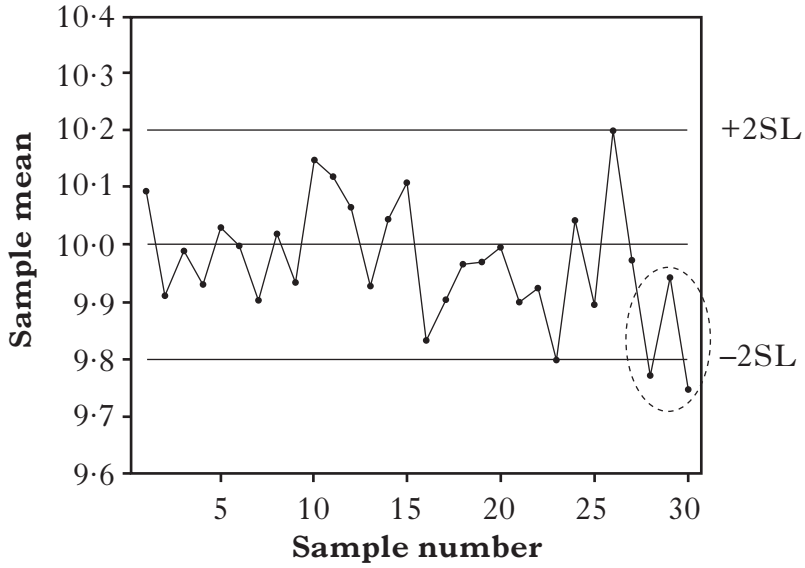
Marks

A8. A manufacturer of automotive parts monitors a critical dimension, X(mm), of a turbo compressor housing using a control chart for the means of samples of size 4, taken from the production line every 30 minutes. When the manufacturing process is unaffected by special cause variation, ie in a state of statistical control, the random variable X has mean 10 and standard deviation 0·2. A chart, with 3-sigma limits, of 30 consecutive samples is shown.



(*a*) Confirm the values of the limits shown and, on the assumptions that X is normally distributed and in a state of statistical control, demonstrate that the tables provided give probability 0·0026 that a sample mean will plot outwith them. **4**

In addition to taking the occurrence of a point outwith 3-sigma limits providing evidence of the presence of special cause variation affecting the process, some control chart practitioners take the occurrence of two out of three consecutive points falling outwith 2-sigma limits on the same side of the centre line to provide such evidence. The second chart shown indicates that this occurs with the final three samples for the 30 samples considered.



A8. (continued)

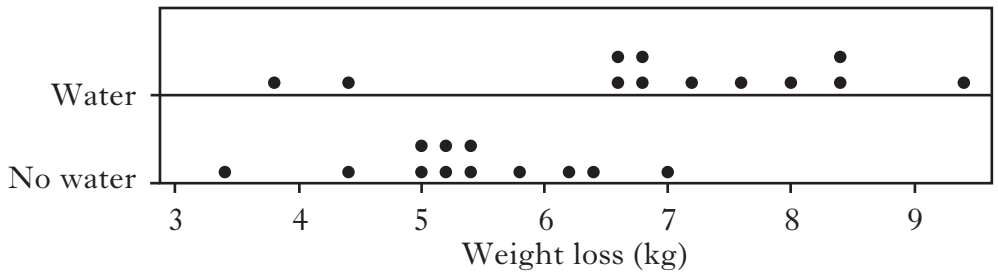
(*b*) Explain why the probability of this type of signal occurring when the process is in a state of statistical control is given by

$$2 \times \binom{3}{2} (0{\cdot}0228)^2 \, (0{\cdot}9772)$$

and show that it has a similar value to that of a point falling outwith 3-sigma limits when the process is in a state of statistical control. **4**

(*c*) Write down and evaluate an expression giving the probability of four out of five consecutive points lying outwith 1-sigma limits on the same side of the centre line, when the process is in statistical control. **3**

A9. In a study of the effect of drinking water prior to meals on weight loss, a sample of 24 subjects with similar characteristics was split into two groups of 12 and each group followed the same diet regime for a period of 10 weeks, the only difference being that subjects in one group were required to drink half a litre of water prior to each meal. A dot plot of the weight losses is displayed.



(*a*) Test the hypothesis that drinking water prior to a meal aids weight loss. **7**

(*b*) A statistical software package gave an approximate 95% confidence interval for the difference in medians to be (0·7, 2·9). Explain how this information is in accord with your work above and state the conclusion of the study in non-statistical terms. **3**

[*END OF SECTION A*]

[Turn over for Section B on *Pages eight and nine*

Marks

Section B (Mathematics for Applied Mathematics)

Answer all the questions

B1. Write down and simplify the general term in the expansion of

$(x^2 + 3x)^8$. **3**

Hence, or otherwise, obtain the coefficient of x^{13}. **2**

B2. (*a*) Given the curve $y = \dfrac{x}{x^2 + 4}$, calculate the gradient when $x = 2$. **3**

(*b*) Determine $\int e^{-2t}\, dt$. **2**

B3. Given $M = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & \lambda \end{pmatrix}$.

(*a*) Calculate M^2. **2**

(*b*) Calculate $M + M^2 + M^3$. **2**

(*c*) For what values of λ does M have an inverse? **2**

B4. Express $\dfrac{1}{x^2 + x}$ in partial fractions, where x is neither 0 nor -1. **3**

A region is enclosed by the curve with equation $y = \dfrac{1}{\sqrt{x^2 + x}}$, the x-axis and the lines $x = 1$ and $x = 3$.

Calculate the volume of the solid of revolution formed by rotating this region through 360° about the x-axis. **4**

Marks

B5. A turkey is taken from a refrigerator to be cooked. Its temperature is 4 °C when it is placed in an oven preheated to 180 °C.

Its temperature, T °C, after a time of x hours in the oven satisfies the equation

$$\frac{dT}{dx} = k(180 - T).$$

(*a*) Show that $T = 180 - 176e^{-kx}$. **4**

(*b*) After an hour in the oven the temperature of the turkey is 30 °C.

Calculate the value of k correct to 2 decimal places. **2**

(*c*) The turkey will be cooked when it reaches a temperature of 80 °C.

After how long (to the nearest minute) will the turkey be cooked? **3**

[*END OF SECTION B*]

[*END OF QUESTION PAPER*]

[BLANK PAGE]

[BLANK PAGE]

[BLANK PAGE]