Topic 7 Part 2 *[433 marks]*

1a. *[3 marks]*

Markscheme

$\mathrm{P}(X \leqslant n) = \sum\limits_{\mathrm{i}=1}^{n} \mathrm{P}(X = \mathrm{i}) = \sum\limits_{\mathrm{i}=1}^{n} pq^{\mathrm{i}-1}$ ***M1A1***

$= p\frac{1-q^n}{1-q}$ ***A1***

$= 1 - (1-p)^n$ ***AG***

[3 marks]

Examiners report

In part (a) some candidates thought that the geometric distribution was continuous, so attempted to integrate the pdf! Others, less seriously, got the end points of the summation wrong.
In part (b) It was very disappointing that may candidates, who got an incorrect answer to part (a), persisted with their incorrect answer into this part.

1b. *[1 mark]*

Markscheme

$(1-p)^m - (1-p)^n$ ***A1***

[1 mark]

Examiners report

In part (a) some candidates thought that the geometric distribution was continuous, so attempted to integrate the pdf! Others, less seriously, got the end points of the summation wrong.
In part (b) It was very disappointing that may candidates, who got an incorrect answer to part (a), persisted with their incorrect answer into this part.

1c. *[2 marks]*

Markscheme

attempt to solve

$0.8 - (0.8)^n > 0.5$ ***M1***

obtain $n = 6$ ***A1***

[2 marks]

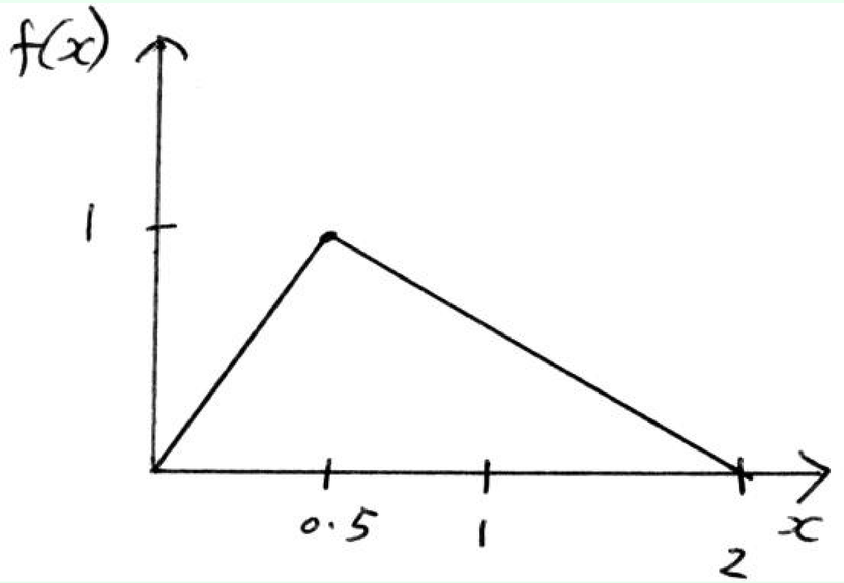
Examiners report

In part (a) some candidates thought that the geometric distribution was continuous, so attempted to integrate the pdf! Others, less seriously, got the end points of the summation wrong.
In part (b) It was very disappointing that may candidates, who got an incorrect answer to part (a), persisted with their incorrect answer into this part.

2a. *[3 marks]*

Markscheme

piecewise linear graph



correct shape ***A1***
with vertices (0, 0), (0.5, 1) and (2, 0) ***A1***
LQ: $x = 0.5$, because the area of the triangle is 0.25 ***R1***
[3 marks]

Examiners report

There was a curious issue about the lower quartile in part (a): The LQ coincides with a quarter of the range of the distribution $\frac{2}{4} = 0.5$. Sadly this is wrong reasoning – the correct reasoning involves a consideration of areas.

2b. *[4 marks]*

Markscheme

(i)
$\mathrm{E}(X) = \int_0^{0.5} x \times 2x\mathrm{d}x + \int_{0.5}^{2} x \times \left(\frac{4}{3} - \frac{2}{3}x\right) \mathrm{d}x = \frac{5}{6}\ (= 0.833...)$ ***(M1)A1***

(ii)
$\mathrm{E}(X^2) = \int_0^{0.5} x^2 \times 2x\mathrm{d}x + \int_{0.5}^{2} x^2 \times \left(\frac{4}{3} - \frac{2}{3}x\right) \mathrm{d}x = \frac{7}{8}\ (= 0.875)$ ***(M1)A1***
[4 marks]

Examiners report

There was a curious issue about the lower quartile in part (a): The LQ coincides with a quarter of the range of the distribution $\frac{2}{4} = 0.5$. Sadly this is wrong reasoning – the correct reasoning involves a consideration of areas.
In part (b) many candidates used hand calculation rather than their GDC.
The random variable Y was not well understood, and that followed into incorrect calculations involving Y – 2X.

2c. *[5 marks]*

Markscheme

(i)
$\mathrm{E}(Y - 2X) = 2\mathrm{E}(X) - 2\mathrm{E}(X) = 0$ ***A1***

(ii)
$\mathrm{Var}\,(X) = \left(\mathrm{E}(X^2) - \mathrm{E}(X)^2\right) = \frac{13}{72}$ ***A1***
$Y = X_1 + X_2 \Rightarrow \mathrm{Var}\,(Y) = 2\mathrm{Var}\,(X)$ ***(M1)***
$\mathrm{Var}\,(Y - 2X) = 2\mathrm{Var}\,(X) + 4\mathrm{Var}\,(X) = \frac{13}{12}$ ***M1A1***
[5 marks]

Examiners report

There was a curious issue about the lower quartile in part (a): The LQ coincides with a quarter of the range of the distribution $\frac{2}{4} = 0.5$. Sadly this is wrong reasoning – the correct reasoning involves a consideration of areas.
In part (b) many candidates used hand calculation rather than their GDC.
The random variable Y was not well understood, and that followed into incorrect calculations involving Y – 2X.

2d. *[7 marks]*

Markscheme

(i) attempt to use
$cf(x) = \int f(u)\mathrm{d}u$ ***M1***

obtain
$$cf(x) = \begin{cases} x^2, & 0 \leqslant x \leqslant 0.5, \\ \frac{4x}{3} - \frac{1}{3}x^2 - \frac{1}{3}, & 0.5 \leqslant x \leqslant 2, \end{cases}$$
$A1$
$A2$

(ii) attempt to solve
$cf(x) = 0.5$ ***M1***
$\frac{4x}{3} - \frac{1}{3}x^2 - \frac{1}{3} = 0.5$ ***(A1)***
obtain 0.775 ***A1***

Note: Accept attempts in the form of an integral with upper limit the unknown median.

Note: Accept exact answer
$2 - \sqrt{1.5}$.

[7 marks]

Examiners report

There was a curious issue about the lower quartile in part (a): The LQ coincides with a quarter of the range of the distribution $\frac{2}{4} = 0.5$. Sadly this is wrong reasoning – the correct reasoning involves a consideration of areas.
In part (b) many candidates used hand calculation rather than their GDC.
The random variable Y was not well understood, and that followed into incorrect calculations involving Y – 2X.

3a. *[2 marks]*

Markscheme

$z = \frac{200-205}{10} = -0.5$ ***(M1)***

probability = 0.691 (accept 0.692) ***A1***

Note: Award ***M1A0*** for 0.309 or 0.308

[2 marks]

Examiners report

As might be expected, (a) was well answered by many candidates, although those who gave 0.6915 straight from tables were given an arithmetic penalty. Parts (b) and (c), however, were not so well answered with errors in calculating the variances being the most common source of incorrect solutions. In particular, some candidates are still uncertain about the difference between nX and $\sum\limits_{i=1}^{n} X_i$.

3b.

[4 marks]

Markscheme

let X be the total weight of the 5 oranges

then

$\mathrm{E}(X) = 5 \times 205 = 1025$ ***(A1)***

$\mathrm{Var}(X) = 5 \times 100 = 500$ ***(M1)(A1)***

$\mathrm{P}(X < 1000) = 0.132$ ***A1***

[4 marks]

Examiners report

As might be expected, (a) was well answered by many candidates, although those who gave 0.6915 straight from tables were given an arithmetic penalty. Parts (b) and (c), however, were not so well answered with errors in calculating the variances being the most common source of incorrect solutions. In particular, some candidates are still uncertain about the difference between nX and $\sum_{i=1}^{n} X_i$.

3c.

[5 marks]

Markscheme

let $Y = B - 3C$ where B is the weight of a random orange and C the weight of a random lemon ***(M1)***

$\mathrm{E}(Y) = 205 - 3 \times 75 = -20$ ***(A1)***

$\mathrm{Var}(Y) = 100 + 9 \times 9 = 181$ ***(M1)(A1)***

$\mathrm{P}(Y > 0) = 0.0686$ ***A1***

[5 marks]

Note: Award ***A1*** for 0.0681 obtained from tables

Examiners report

As might be expected, (a) was well answered by many candidates, although those who gave 0.6915 straight from tables were given an arithmetic penalty. Parts (b) and (c), however, were not so well answered with errors in calculating the variances being the most common source of incorrect solutions. In particular, some candidates are still uncertain about the difference between nX and $\sum_{i=1}^{n} X_i$.

4a. *[5 marks]*

Markscheme

the weight losses are

2.2

3.5

4.3

–0.5

4.2

–0.2

2.5

2.7

0.1

–0.7 ***(M1)(A1)***

$\sum x = 18.1$,

$\sum x^2 = 67.55$

UE of mean = 1.81 ***A1***

UE of variance

$= \frac{67.55}{9} - \frac{18.1^2}{90} = 3.87$ ***(M1)A1***

Note: Accept weight losses as positive or negative. Accept unbiased estimate of mean as positive or negative.

Note: Award ***M1A0*** for 1.97 as UE of variance.

[5 marks]

Examiners report

In (a), most candidates gave a correct estimate for the mean but the variance estimate was often incorrect. Some candidates who use their GDC seem to be unable to obtain the unbiased variance estimate from the numbers on the screen. The way to proceed, of course, is to realise that the larger of the two 'standard deviations' on offer is the square root of the unbiased estimate so that its square gives the required result. In (b), most candidates realised that the t-distribution should be used although many were awarded an arithmetic penalty for giving either $t = 2.911$ or the critical value = 2.821. Some candidates who used the p-value method to reach a conclusion lost a mark by omitting to give the critical value. Many candidates found part (c) difficult and although they were able to obtain $t = 2.49\ldots$, they were then unable to continue to obtain the confidence interval.

4b.

[6 marks]

Markscheme

(i)

$H_0 : \mu_d = 0$ versus

$H_1 : \mu_d > 0$ ***A1***

Note: Accept any symbol for

μ_d

(ii) using *t* test ***(M1)***

$t = \frac{1.81}{\sqrt{\frac{3.87}{10}}} = 2.91$ ***A1***

(iii) DF = 9 ***(A1)***

Note: Award this ***(A1)*** if the p-value is given as 0.00864

1% critical value = 2.82 ***A1***

accept

H_1 ***R1***

Note: Allow ***FT*** on final ***R1***.

[6 marks]

Examiners report

In (a), most candidates gave a correct estimate for the mean but the variance estimate was often incorrect. Some candidates who use their GDC seem to be unable to obtain the unbiased variance estimate from the numbers on the screen. The way to proceed, of course, is to realise that the larger of the two 'standard deviations' on offer is the square root of the unbiased estimate so that its square gives the required result. In (b), most candidates realised that the t-distribution should be used although many were awarded an arithmetic penalty for giving either $t = 2.911$ or the critical value = 2.821. Some candidates who used the *p*-value method to reach a conclusion lost a mark by omitting to give the critical value. Many candidates found part (c) difficult and although they were able to obtain $t = 2.49\ldots$, they were then unable to continue to obtain the confidence interval.

5a. *[5 marks]*

Markscheme

under

H_0 ,

S is $\mathrm{Po}(30)$ ***(A1)***

EITHER

$\mathrm{P}(S \leqslant 22) = 0.080569\ldots$ ***A1***

$\mathrm{P}(S \geqslant 38) = 0.089012\ldots$ ***A1***

significance level = 0.080569… + 0.089012… ***(M1)***

= 0.170 ***A1***

OR

$\mathrm{P}(S \leqslant 22) = 0.080569\ldots$ ***A1***

$\mathrm{P}(S \leqslant 37) = 0.910987\ldots$ ***A1***

significance level = 1 – (0.910987…) + 0.089012… ***(M1)***

= 0.170 ***A1***

Note: Accept 17 % or 0.17.

Note: Award 2 marks out of the final 4 marks for correct use of the Central Limit Theorem, giving 0.144 without a continuity correction and 0.171 with a continuity correction. The first ***(A1)*** is independent.

[5 marks]

Examiners report

Solutions to this question were often disappointing with many candidates not knowing what had to be done. Even those candidates who knew what to do sometimes made errors in evaluating the probabilities, often by misinterpreting the inequality signs. Candidates who used the Central Limit Theorem to evaluate the probabilities were given only partial credit on the grounds that the answers obtained were approximate and not exact.

5b.

[5 marks]

Markscheme

S is now Po(25)   ***(A1)***

P (Type II error) = P (accept

$H_0|\mu = 2.5)$ ***(M1)***

$= \mathrm{P}\left(23 \leqslant S \leqslant 37|S \text{ is Po}(25)\right)$ ***(M1)***

Note: Only one of the above ***M1*** marks can be implied.

= 0.990789… – 0.317533…   ***(A1)***

= 0.673   ***A1***

Note: Award 2 marks out of the final 4 marks for correct use of the Central Limit Theorem, giving 0.647 without a continuity correction and 0.685 with a continuity correction. The first ***(A1)*** is independent.

[5 marks]

Examiners report

Solutions to this question were often disappointing with many candidates not knowing what had to be done. Even those candidates who knew what to do sometimes made errors in evaluating the probabilities, often by misinterpreting the inequality signs. Candidates who used the Central Limit Theorem to evaluate the probabilities were given only partial credit on the grounds that the answers obtained were approximate and not exact.

6a.

[7 marks]

Markscheme

(i)

$f(x) = \binom{x-1}{2} p^3 (1-p)^{x-3}$ ***M1A1***

Note: Award ***M1A0*** for

$f(x) = \binom{x-1}{2} p^3 q^{x-3}$

taking logs, ***M1***

$\ln f(x) = \left(\ln \binom{x-1}{2} p^3 (1-p)^{x-3}\right)$

$= \ln\left(\frac{(x-1)(x-2)}{2} \times p^3 (1-p)^{x-3}\right)$ ***A1***

Note: Award ***A1*** for simplifying binomial coefficient, seen anywhere.

$= \ln\left(\frac{(x-1)(x-2)}{2} \times p^3 \frac{(1-p)^x}{(1-p)^3}\right)$ ***A1***

Note: Award ***A1*** for correctly splitting

$(1-p)^{x-3}$, seen anywhere.

$= 3\ln\left(\frac{p}{1-p}\right) + \ln(x-1) + \ln(x-2) + x\ln(1-p) - \ln 2$ ***AG***

(ii) the domain is $\{3, 4, 5, \ldots\}$ ***A1***

Note: Do not accept

$x \geqslant 3$

(iii) differentiating with respect to x , ***M1***

$\frac{f'(x)}{f(x)} = \frac{1}{x-1} + \frac{1}{x-2} + \ln(1-p)$ ***AG***

[7 marks]

Examiners report

In general, candidates were able to start this question, but very few wholly correct answers were seen. Most candidates were able to write down the probability function but the process of taking logs was often unconvincing. The vast majority of candidates gave an incorrect domain for *f*, the most common error being

$x \geqslant 3$. Most candidates failed to realise that the solution to (b) was to be found by setting the right-hand side of the given equation equal to zero. Many of the candidates who obtained the correct answer, 6.195…, then rounded this to 6 without realising that both 6 and 7 should be checked to see which gave the larger probability.

6b. *[5 marks]*

Markscheme

setting

$f'(x) = 0$ and putting $p = 0.35$,

$\frac{1}{x-1} + \frac{1}{x-2} + \ln 0.65 = 0$ ***M1A1***

solving, $x = 6.195\ldots$ ***A1***

we need to check $x = 6$ and 7

$f(6) = 0.1177\ldots$ and $f(7) = 0.1148\ldots$ ***A1***

the most likely value of Y is 6 ***A1***

Note: Award the final ***A1*** for the correct conclusion even if the previous ***A1*** was not awarded.

[5 marks]

Examiners report

In general, candidates were able to start this question, but very few wholly correct answers were seen. Most candidates were able to write down the probability function but the process of taking logs was often unconvincing. The vast majority of candidates gave an incorrect domain for *f*, the most common error being $x \geqslant 3$. Most candidates failed to realise that the solution to (b) was to be found by setting the right-hand side of the given equation equal to zero. Many of the candidates who obtained the correct answer, 6.195…, then rounded this to 6 without realising that both 6 and 7 should be checked to see which gave the larger probability.

7. *[13 marks]*

Markscheme

(a)

$H_0 : \mu = 2.5$ ***A1***

$H_1 : \mu \neq 2.5$ ***A1***

[2 marks]

(b) the critical values are

$2.5 \pm 1.96 \times \frac{0.1}{\sqrt{16}}$, ***(M1)(A1)(A1)***

i.e. 2.45, 2.55 ***(A1)***

the critical region is

$\bar{x} < 2.45 \cup \bar{x} > 2.55$ ***A1A1***

Note: Accept

$\leqslant$, $\geqslant$.

[6 marks]

(c)

$\bar{X}$ is now

$\mathrm{N}(2.6,\ 0.025^2)$ ***A1***

a Type II error is accepting

H_0 when

H_1 is true ***(R1)***

thus we require

$\mathrm{P}(2.45 < \bar{X} < 2.55)$ ***M1A1***

$= 0.0228$ (Accept 0.0227) ***A1***

Note: If critical values of 2.451 and 2.549 are used, accept 0.0207.

[5 marks]

Total [13 marks]

Examiners report

In (a), some candidates incorrectly gave the hypotheses in terms of

$\bar{x}$ instead of

μ. In (b), many candidates found the correct critical values but then some gave the critical region as

$2.45 < \bar{x} < 2.55$ instead of

$\bar{x} < 2.45 \cup \bar{x} > 2.55$ Many candidates gave the critical values correct to four significant figures and therefore were given an arithmetic penalty. In (c), many candidates correctly defined a Type II error but were unable to calculate the corresponding probability.

8a. *[9 marks]*

Markscheme

let

$\bar{A}$, $\bar{B}$ denote the means of Alan's and Brian's jumps

attempting to find the distributions of

$\bar{A}$, $\bar{B}$ ***(M1)***

$\bar{A}$ is $\mathrm{N}\left(5.2, \frac{0.1^2}{4}\right)$ ***A1***

$\bar{B}$ is $\mathrm{N}\left(5.1, \frac{0.12^2}{3}\right)$ ***A1***

attempting to find the distribution of

$\bar{A} - \bar{B}$ ***(M1)***

$\bar{A} - \bar{B}$ is $\mathrm{N}\left(5.2 - 5.1, \frac{0.1^2}{4} + \frac{0.12^2}{3}\right)$ ***(A1)(A1)***

i.e.

$\mathrm{N}(0.1,\ 0.0073)$ ***A1***

$\mathrm{P}(\bar{A} < \bar{B}) = \mathrm{P}(\bar{A} - \bar{B} < 0)$ ***M1***

$= 0.121$ ***A1***

[9 marks]

Examiners report

In (a), it was disappointing to note that many candidates failed to realise that the question was concerned with the mean lengths of the jumps and worked instead with the sums of the lengths.

8b. *[10 marks]*

$\sum x = 31.38,\ \sum x^2 = 164.1294$

$\bar{x} = \frac{31.38}{6} = 5.23$

$s_{n-1}^2 = \frac{164.1294}{5} - \frac{31.38^2}{5 \times 6} = 0.00240$

$s_{n-1} = 0.04899 \Rightarrow s_{n-1}^2 = 0.00240$

$5.23 \pm 2.015\sqrt{\frac{0.0024}{6}}$

Examiners report

Most candidates obtained correct estimates in (b)(i), usually directly from the GDC. In (b)(ii), however, some candidates found a *z*-interval instead of a *t*-interval.

9. *[12 marks]*

Markscheme

(a) estimate of

$\mu = 13.1$ ***A1***

estimate of

$\sigma^2 = 0.416$ ***A1***

[2 marks]

(b) using a GDC (or otherwise), the 95% confidence interval is ***(M1)***

[12.6, 13.6] ***A1A1***

Note: Accept open or closed intervals.

[3 marks]

(c) (i)

$t = \frac{13.1-12.5}{0.6446.../\sqrt{10}} = 2.94$ ***(M1)A1***

$v = 9$ ***(A1)***

p-value

$= 2 \times \mathrm{P}(T > 2.9433\ldots)$ ***(M1)***

$= 0.0164$ (accept 0.0165) ***A1***

(ii) we accept the null hypothesis (the mean travel time is 12.5 minutes) ***A1***

because 0.0164 (or 0.0165) > 0.01 ***R1***

Note: Allow follow through on their *p*-value.

[7 marks]

Total [12 marks]

Examiners report

This was well answered by many candidates. In (a), some candidates chose the wrong standard deviation from their calculator and often failed to square their result to obtain the unbiased variance estimate. Candidates should realise that it is the smaller of the two values (ie the one obtained by dividing by (n – 1)) that is required. The most common error was to use the normal distribution instead of the *t*-distribution. The signpost towards the *t*-distribution is the fact that the variance had to be estimated in (a). Accuracy penalties were often given for failure to round the confidence limits, the *t*-statistic or the *p*-value to three significant figures.

10.

[10 marks]

Markscheme

(a) let

$T = \sum_{i=1}^{10} X_i$ so that T is Po(10) under

H_0 ***(M1)***

$\mathrm{P(Type\ I\ error)} = \mathrm{P}\ T \geqslant 15 | \mu = 1$ ***M1A1***

$= 0.0835$ ***A2*** ***N3***

Note: Candidates who write the first line and only the correct answer award ***(M1)M0A0A2***.

[5 marks]

(b) let

$T = \sum_{i=1}^{10} X_i$ so that T is Po(20) under

H_1 ***(M1)***

$\mathrm{P(Type\ II\ error)} = \mathrm{P}\ T \leqslant 14 | \mu = 2$ ***M1A1***

$= 0.105$ ***A2*** ***N3***

Note: Candidates who write the first line and only the correct answer award ***(M1)M0A0A2***.

Note: Award 5 marks to a candidate who confuses Type I and Type II errors and has both answers correct.

[5 marks]

Total [10 marks]

Examiners report

This question caused problems for many candidates and the solutions were often disappointing. Some candidates seemed to be unaware of the meaning of Type I and Type II errors. Others were unable to calculate the probabilities even when they knew what they represented. Candidates who used a normal approximation to obtain the probabilities were not given full credit – there seems little point in using an approximation when the exact value could be found.

11.

[8 marks]

Markscheme

let X, Y, Z denote respectively the weights, in grams, of a randomly chosen apple, pear, peach

then

$U = X + Y - 2Z$ is $\mathrm{N}(115 + 110 - 2 \times 105,\ 5^2 + 4^2 + 2^2 \times 3^2)$ ***(M1)(A1)(A1)***

Note: Award ***M1*** for attempted use of U.

i.e. N(15, 77) ***A1***

we require

$\mathrm{P}(X + Y > 2Z) = \mathrm{P}(U > 0)$ ***M1A1***

$= 0.956$ ***A2***

Note: Award ***M0A0A2*** for 0.956 only.

[8 marks]

Examiners report

Solutions to this question again illustrated the fact that many candidates are unable to distinguish between nX and $\sum_{i=1}^{n} X_i$ so that many candidates obtained an incorrect variance to evaluate the final probability.

12.

[10 marks]

Markscheme

$\mathrm{P}(X=10)=\binom{9}{4}p^5(1-p)^5 \quad (=0.05)$ ***(M1)A1A1***

Note: First ***A1*** is for the binomial coefficient. Second ***A1*** is for the rest.

solving by any method,

$p=0.297\ldots$ ***A4***

Notes: Award ***A2*** for anything which rounds to 0.703.

Do not apply any ***AP*** at this stage.

$\mathrm{P}(X=10)=\binom{10}{4}\times(0.297...)^5\times(1-0.297...)^6$ ***(M1)A1***

= 0.0586 ***A1***

Note: Allow follow through for incorrect p-values.

[10 marks]

Examiners report

Questions on these discrete distributions have not been generally well answered in the past and it was pleasing to note that many candidates submitted a reasonably good solution to this question. In (b), the determination of the value of p was often successful using a variety of methods including solving the equation

$p(1-p)=(0.000396\ \ldots)^{1/5}$, graph plotting or using SOLVER on the GDC or even expanding the equation into a 10^{th} degree polynomial and solving that. Solutions to this particular question exceeded expectations.

13.

[4 marks]

Markscheme

$\bar{m}=\frac{6.7+7.2+\ldots+7.3}{10}=6.91$ ***(M1)A1***

$s_{n-1}^2=\frac{1}{9}\left((6.7-6.91)^2+\ldots+(7.3-6.91)^2\right)$ ***(M1)***

$=\frac{0.489}{9}=0.0543\text{ (3 sf)}$ ***A1***

Note: Award ***M1A0*** for 0.233.

[4 marks]

Examiners report

Most candidates used a GDC to answer this question and many scored full marks in this question. However there were a significant number of candidates who showed little understanding of the meaning of unbiased estimate. In some cases, candidates wasted time by attempting to calculate the required values by hand.

14.

[7 marks]

Markscheme

let

$W = \sum_{i=1}^{6} w_i$ ***(M1)***

w_i is $\mathrm{N}(200,\ 15^2)$

$\mathrm{E}(W) = \sum_{i=1}^{6} \mathrm{E}(w_i) = 6 \times 200 = 1200$ ***A1***

$\mathrm{Var}(W) = \sum_{i=1}^{6} \mathrm{Var}(w_i) = 6 \times 15^2 = 1350$ ***A2***

W is $\mathrm{N}(1200,\ 1350)$ ***(M1)***

$\mathrm{P}(W > 1150) = 0.913$ by GDC ***A1A1***

Note: Using 6 times the mean or a lower bound for the mean are acceptable methods.

[7 marks]

Examiners report

Candidates will often be asked to solve these problems that test if they can distinguish between a number of individuals and a number of copies. The wording of the question was designed to make the difference clear. If candidates wrote

$w_1 + \ldots + w_6$ in (a) and $12w$ in (b), they usually went on to gain full marks.

15.

[15 marks]

Markscheme

(a) (i) as

σ^2 is known

$\bar{x}$ is

$\mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$ ***(M1)***

CI is

$\bar{x} - z^* \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z^* \frac{\sigma}{\sqrt{n}}$ ***(M1)***

$\bar{x} = 9.24,\ z^* = 1.960$ for 95 % CI ***(A1)***

CI is

$6.61 < \mu < 11.87$ by GDC ***A1A1***

(ii) CI is

$\bar{x} - z^* \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z^* \frac{\sigma}{\sqrt{n}}$

require

$2 \times 1.645 \frac{3}{\sqrt{n}} < 2$ ***R1A1***

$4.935 < \sqrt{n}$ ***(A1)***

$24.35 < n$ ***A1***

so smallest value for $n = 25$ ***A1***

Note: Accept use of table.

[10 marks]

(b) as

σ^2 is not known

$\bar{x}$ has the t distribution with $v = 4$ ***(M1)(A1)***

CI is

$\bar{x} - t^* \frac{s_{n-1}}{\sqrt{n}} < \mu < \bar{x} + t^* \frac{s_{n-1}}{\sqrt{n}}$

$\bar{x} = 9.24,\ s_{n-1} = 5.984,\ t^* = 2.776$ for 95 % CI ***(A1)***

CI is

$1.81 < \mu < 16.67$ by GDC ***A1A1***

[5 marks]

Total [15 marks]

Examiners report

The 2 confidence intervals were generally done well by using a calculator. Some marks were dropped by not giving the answers to 2 decimal places as required. Weak candidates did not realise that (b) was a t interval. Part (a) (ii) was not as well answered and often it was the first step that was the problem.

16.

[10 marks]

Markscheme

(a) we are dealing with the Negative Binomial distribution:

$\mathrm{NB}\left(4, \frac{1}{3}\right)$ ***(M1)***

let X be the number of scheduled lessons before the email is sent

$\mathrm{P}(X=8)=\binom{7}{3}\left(\frac{2}{3}\right)^4\left(\frac{1}{3}\right)^4=0.0854$ ***(M1)A1***

[3 marks]

(b)

$\mathrm{E}(X)=\frac{r}{p}=\frac{4}{\frac{1}{3}}=12$ ***(M1)A1***

[2 marks]

(c) we are asking for 2 missed lessons in the second 6 lessons, with the last lesson missed so this is

$\mathrm{NB}\left(2, \frac{1}{3}\right)$ ***(M1)***

$\mathrm{P}(X=6)=\binom{5}{1}\left(\frac{2}{3}\right)^4\left(\frac{1}{3}\right)^2=0.110$ ***(M1)A1***

Note: Accept solutions laid out in terms of conditional probabilities.

[3 marks]

(d) **EITHER**

We know that she missed the

6^{th} lesson so she must have missed 3 from the first 5 lessons. All are equally likely so the probability that she missed the

2^{nd} lesson is

$\frac{3}{5}$. ***R1A1***

OR

require

$\mathrm{P}(\text{missed } 2^{\mathrm{nd}} \mid X=6)=\frac{\mathrm{P}(\text{missed } 2^{\mathrm{nd}} \text{ and } X=6)}{\mathrm{P}(X=6)}$ ***R1***

$\mathrm{P}(\text{missed } 2^{\mathrm{nd}} \text{ and } X=6)=\mathrm{P}(\text{missed } 2^{\mathrm{nd}} \text{ and } 6^{\mathrm{th}} \text{ and 2 of remaining 4})$

$$=\frac{1}{3}\cdot\frac{1}{3}\cdot\binom{4}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^2=\frac{24}{3^6}$$

$\mathrm{P}(X=6)=\binom{5}{3}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^2=\frac{40}{3^6}$

so required probability is

$\frac{24}{3^6}\cdot\frac{3^6}{40}=\frac{3}{5}$ ***A1***

[2 marks]

Total [10 marks]

Examiners report

Realising that this was a problem about the Negative Binomial distribution was the crucial thing to realise in this question. All parts of the syllabus do need to be covered.

17. *[9 marks]*

Markscheme

(a) a Type I error is when

H_0 is rejected, when

H_0 is actually true ***A1***

[1 mark]

(b)

$P(H_0 \text{ rejected}|H_0 \text{ true}) = P(\text{at least one R}|6 \text{ J and } 2 \text{ R})$ ***M1***

EITHER

$P(\text{no R}|H_0 \text{ true}) = \frac{6}{8} \times \frac{5}{7} = \frac{15}{28}$ ***(A1)***

OR

let X count the number of R's given by the computer under

H_0, $X \sim \text{Hyp}(2, 2, 8)$

$P(X = 0) = \frac{\binom{2}{0}\binom{6}{2}}{\binom{8}{2}} = \frac{15}{28}$ ***(A1)***

THEN

$P(\text{at least one R}|H_0 \text{ true}) = 1 - \frac{15}{28}$ ***(M1)***

$P(\text{Type I error}) = \frac{13}{28}$ $(= 0.464)$ ***A1***

[4 marks]

(c) a Type II error is when

H_0 is accepted, when

H_0 is actually false ***A1***

[1 mark]

(d)

$P(H_0 \text{ accepted}|H_0 \text{ false}) = P(2 \text{ J}|3 \text{ J and } 4 \text{ R})$ ***M1***

EITHER

$P(2 \text{ J}|H_0 \text{ false}) = \frac{3}{7} \times \frac{2}{6} = \frac{1}{7}$ ***(A1)***

OR

let Y count the number of R's given by the computer.

H_0 false implies

$Y \sim \text{Hyp}(2, 4, 7)$

$P(Y = 0) = \frac{\binom{4}{0}\binom{3}{2}}{\binom{7}{2}} = \frac{1}{7}$ ***(A1)***

THEN

$P(\text{Type II error}) = \frac{1}{7} (= 0.143)$ ***A1***

[3 marks]

Total [9 marks]

Examiners report

Poorer candidates just gained the 2 marks for saying what a Type I and Type II error were and could not then apply the definitions to obtain the conditional probabilities required. It was clear from some crossings out that even the 2 definition continue to cause confusion. Good, clear-thinking candidates were able to do the question correctly.

18a. *[2 marks]*

Markscheme

let S be the weight of tea in a random *Supermug* tea bag

$S \sim \mathrm{N}(4.2,\, 0.15^2)$

$\mathrm{P}(S > 3.9) = 0.977$ ***(M1)A1***

[2 marks]

Examiners report

For most candidates this was a reasonable start to the paper with many candidates gaining close to full marks. The most common error was in (b) where, surprisingly, many candidates did not realise the need to multiply the product of the two probabilities by 2 to gain the final answer. Weaker candidates often found problems in understanding how to correctly find the variance in both (c) and (d).

18b. *[4 marks]*

Markscheme

let M be the weight of tea in a random *Megamug* tea bag

$M \sim \mathrm{N}(5.6,\, 0.17^2)$

$\mathrm{P}(M > 5.4) = 0.880\ldots$ ***(A1)***

$\mathrm{P}(M < 5.4) = 1 - 0.880\ldots = 0.119\ldots$ ***(A1)***

required probability

$= 2 \times 0.880\ldots \times 0.119\ldots = 0.211$ ***M1A1***

[4 marks]

Examiners report

For most candidates this was a reasonable start to the paper with many candidates gaining close to full marks. The most common error was in (b) where, surprisingly, many candidates did not realise the need to multiply the product of the two probabilities by 2 to gain the final answer. Weaker candidates often found problems in understanding how to correctly find the variance in both (c) and (d).

18c. *[4 marks]*

Markscheme

$\mathrm{P}(S_1 + S_2 + S_3 + S_4 + S_5 < 20.5)$

let

$S_1 + S_2 + S_3 + S_4 + S_5 = A$ ***(M1)***

$\mathrm{E}(A) = 5\mathrm{E}(S)$

= 21 ***A1***

$\mathrm{Var}(A) = 5\mathrm{Var}(S)$

= 0.1125 ***A1***

$A \sim \mathrm{N}(21,\, 0.1125)$

$\mathrm{P}(A < 20.5) = 0.0680$ ***A1***

[4 marks]

Examiners report

For most candidates this was a reasonable start to the paper with many candidates gaining close to full marks. The most common error was in (b) where, surprisingly, many candidates did not realise the need to multiply the product of the two probabilities by 2 to gain the final answer. Weaker candidates often found problems in understanding how to correctly find the variance in both (c) and (d).

18d. *[5 marks]*

Markscheme

$\mathrm{P}(S_1 + S_2 + S_3 + S_4 + S_5 + S_6 + S_7 - (M_1 + M_2 + M_3 + M_4 + M_5) > 0)$

let

$S_1 + S_2 + S_3 + S_4 + S_5 + S_6 + S_7 - (M_1 + M_2 + M_3 + M_4 + M_5) = B$ ***(M1)***

$\mathrm{E}(B) = 7\mathrm{E}(S) - 5\mathrm{E}(M)$

= 1.4 ***A1***

Note: Above ***A1*** is independent of first ***M1***.

$\mathrm{Var}(B) = 7\mathrm{Var}(S) + 5\mathrm{Var}(M)$ ***(M1)***

= 0.302 ***A1***

$\mathrm{P}(B > 0) = 0.995$ ***A1***

[5 marks]

Examiners report

For most candidates this was a reasonable start to the paper with many candidates gaining close to full marks. The most common error was in (b) where, surprisingly, many candidates did not realise the need to multiply the product of the two probabilities by 2 to gain the final answer. Weaker candidates often found problems in understanding how to correctly find the variance in both (c) and (d).

19a. *[3 marks]*

Markscheme

$\int \lambda e^{-\lambda t} dt = -e^{-\lambda t}\ (+c)$ ***A1***

$\Rightarrow F(x) = \left[-e^{-\lambda t}\right]_0^x$ ***(M1)***

$= 1 - e^{-\lambda t}\ (x \geqslant 0)$ ***A1***

[3 marks]

Examiners report

For most candidates the question started well, but many did not appear to understand how to find the cumulative distribution function in (b). Many were able to integrate
$\lambda e^{-\lambda x}$, but then did not know what to do with the integral. Parts (c), (d) and (e) were relatively well done, but even candidates who successfully found the cumulative distribution function often did not use it. This resulted in a lot of time spent integrating the same function.

19b. *[2 marks]*

Markscheme

$1 - F\left(\frac{2}{\lambda}\right)$ ***M1***

$= e^{-2}\ \ (= 0.135)$ ***A1***

[2 marks]

Examiners report

For most candidates the question started well, but many did not appear to understand how to find the cumulative distribution function in (b). Many were able to integrate
$\lambda e^{-\lambda x}$, but then did not know what to do with the integral. Parts (c), (d) and (e) were relatively well done, but even candidates who successfully found the cumulative distribution function often did not use it. This resulted in a lot of time spent integrating the same function.

19c. *[3 marks]*

Markscheme

$F(m) = \frac{1}{2}$ ***(M1)***

$\Rightarrow e^{-\lambda m} = \frac{1}{2}$ ***A1***

$\Rightarrow -\lambda m = \ln\frac{1}{2}$

$\Rightarrow m = \frac{1}{\lambda}\ln 2$ ***A1***

[3 marks]

Examiners report

For most candidates the question started well, but many did not appear to understand how to find the cumulative distribution function in (b). Many were able to integrate
$\lambda e^{-\lambda x}$, but then did not know what to do with the integral. Parts (c), (d) and (e) were relatively well done, but even candidates who successfully found the cumulative distribution function often did not use it. This resulted in a lot of time spent integrating the same function.

19d. *[2 marks]*

Markscheme

$F\left(\frac{1}{\lambda}\right) - F\left(\frac{\ln 2}{\lambda}\right)$ ***M1***

$= \frac{1}{2} - e^{-1} \quad (= 0.132)$ ***A1***

[2 marks]

Examiners report

For most candidates the question started well, but many did not appear to understand how to find the cumulative distribution function in (b). Many were able to integrate
$\lambda e^{-\lambda x}$, but then did not know what to do with the integral. Parts (c), (d) and (e) were relatively well done, but even candidates who successfully found the cumulative distribution function often did not use it. This resulted in a lot of time spent integrating the same function.

20a. *[2 marks]*

Markscheme

(i) median = 104 grams ***A1***

Note: Accept 105.

(ii) 30^{th} percentile = 90 grams ***A1***

[2 marks]

Examiners report

On a very straightforward question there were many correct answers. However, there was evidence that some candidates had not previously encountered cumulative frequency graphs and hence scored low marks on the question.

20b. *[2 marks]*

Markscheme

$80 - 49$ ***(M1)***

$= 31$ ***A1***

Note: Accept answers 30 to 32.

[2 marks]

Examiners report

On a very straightforward question there were many correct answers. However, there was evidence that some candidates had not previously encountered cumulative frequency graphs and hence scored low marks on the question.

21. *[8 marks]*

Markscheme

The number of' 'heads' X is B(200, p) ***(M1)***

$\mathrm{H}_0 : p = 0.5$; $\mathrm{H}_1 : p \neq 0.5$ ***A1A1***

Note: Award ***A1A0*** for the statement "

H_0 : coin is fair;

H_1 : coin is biased".

EITHER

$\mathrm{P}(X \geqslant 115|\mathrm{H}_0) = 0.0200$ ***(M1)(A1)***

p-value = 0.0400 ***A1***

This is greater than 0.01. ***R1***

There is insufficient evidence to conclude that the coin is biased (or the coin is not biased). ***R1***

OR

(Using a proportion test on a GDC) p-value = 0.0339 ***N3***

This is greater than 0.01. ***R1***

There is insufficient evidence to conclude that the coin is biased (or the coin is not biased). ***R1***

OR

Under

H_0 X is approximately N(100, 50) ***(M1)***

$z = \frac{115-100}{\sqrt{50}} = 2.12$ ***(M1)A1***

(Accept 2.05 with continuity correction)

This is less than 2.58 ***R1***

There is insufficient evidence to conclude that the coin is biased (or the coin is not biased). ***R1***

OR

99 % confidence limits for p are

$\frac{115}{200} \pm 2.576\sqrt{\frac{115}{200} \times \frac{85}{200} \times \frac{1}{200}}$ ***(M1)A1***

giving [0.485, 0.665] ***A1***

This interval contains 0.5 ***R1***

There is insufficient evidence to conclude that the coin is biased (or the coin is not biased). ***R1***

[8 marks]

Examiners report

This question was well answered in general with several correct methods seen. The most popular method was to use a GDC to carry out a proportion test which is equivalent to using a normal approximation. Relatively few candidates calculated an exact p-value using the binomial distribution. Candidates who found a 95% confidence interval for p, the probability of obtaining a head, and noted that this contained 0.5 were given full credit.

22.

[14 marks]

Markscheme

(a) Let X, Y (grams) denote respectively the weights of a randomly chosen apple, pear.

Then

$X - 2Y$ is $\mathrm{N}(200 - 2 \times 120,\ 15^2 + 4 \times 10^2)$, ***(M1)(A1)(A1)***

i.e.

$\mathrm{N}(-40,\ 25^2)$ ***A1***

We require

$\mathrm{P}(X > 2Y) = \mathrm{P}(X - 2Y > 0)$ ***(M1)(A1)***

$= 0.0548$ ***A2***

[8 marks]

(b) Let

$T = X_1 + X_2 + X_3 + Y_1 + Y_2 + Y_3 + Y_4$ (grams) denote the total weight.

Then

T is $\mathrm{N}(3 \times 200 + 4 \times 120,\ 3 \times 15^2 + 4 \times 10^2)$, ***(M1)(A1)(A1)***

i.e.

$\mathrm{N}(1080, 1075)$ ***A1***

$\mathrm{P}(T > 1000) = 0.993$ ***A2***

[6 marks]

Total [14 marks]

Examiners report

The response to this question was disappointing. Many candidates are unable to differentiate between quantities such as $3X$ and $X_1 + X_2 + X_3$. While this has no effect on the mean, there is a significant difference between the variances of these two random variables.

23. *[14 marks]*

Markscheme

(a) (i)
$\bar{x} = \frac{4.35+4.53}{2} = 4.44$ (estimate of
μ) ***A2***

(ii) Degrees of freedom = 9 ***(A1)***
Critical value of *t* = 2.262 ***(A1)***
$2.262 \times \frac{s}{\sqrt{10}} = 0.09$ ***M1A1***
$s = 0.12582\ldots$ ***(A1)***
$s^2 = 0.0158$ (estimate of
σ^2) ***A1***
[8 marks]

(b) (i) Using *t* test ***(M1)***
$t = \frac{4.44-4.5}{\sqrt{\frac{0.0158}{10}}} = -1.50800$ (Accept
-1.50946) ***(A1)***
p-value = 0.0829 (Accept 0.0827) ***A2***

(ii) (a) Accept
H_0 / Reject
H_1 . ***R1***
(b) Reject
H_0 / Accept
H_1 . ***R1***
[6 marks]

Total [14 marks]

Examiners report

Most candidates realised that the unbiased estimate of the mean was simply the central point of the confidence interval. Many candidates, however, failed to realise that, because the variance was unknown, the *t*-distribution was used to determine the confidence limits. In (b), although the *p*-value was asked for specifically, some candidates solved the problem correctly by comparing the value of their statistic with the appropriate critical values. This method was given full credit but, of course, marks were lost by their failure to give the *p*-value.

24. *[10 marks]*

Markscheme

(a)

X_2 is a geometric random variable ***A1***

with

$p = \frac{5}{6}$. ***A1***

Therefore

$\mathrm{E}(X_2) = \frac{6}{5}$. ***A1***

[3 marks]

(b)

X_3 is a geometric random variable with

$p = \frac{4}{6}$. ***A1***

Therefore

$\mathrm{E}(X_3) = \frac{6}{4}$. ***A1***

[2 marks]

(c)

$\mathrm{E}(X_4) = \frac{6}{3}$, $\mathrm{E}(X_5) = \frac{6}{2}$, $\mathrm{E}(X_6) = \frac{6}{1}$ ***A1A1A1***

$\mathrm{E}(X_1) = 1$ (or $X_1 = 1$) ***A1***

Expected number of tosses

$\sum\limits_{n=1}^{6} \mathrm{E}(X_n)$ ***M1***

$= 14.7$ ***AG***

[5 marks]

Total [10 marks]

Examiners report

Many candidates were unable even to start this question although those who did often made substantial progress.

25a. *[6 marks]*

Markscheme

(i)

$\mathrm{E}(2Y + 3) = 6$

$2\mathrm{E}(Y) + 3 = 6$ ***M1***

$\mathrm{E}(Y) = \frac{3}{2}$ ***A1***

(ii)

$\mathrm{Var}(2 - 3Y) = 11$

$\mathrm{Var}(-3Y) = 11$ ***(M1)***

$9\mathrm{Var}(Y) = 11$

$\mathrm{Var}(Y) = \frac{11}{9}$ ***A1***

(iii)

$\mathrm{E}(Y^2) = \mathrm{Var}(Y) + [\mathrm{E}(Y)]^2$ ***M1***

$= \frac{11}{9} + \frac{9}{4}$

$= \frac{125}{36}$ ***A1*** ***N0***

[6 marks]

Examiners report

$E(Y)$ was calculated correctly but many could not go further to find

$Var(Y)$ and $E(Y^2)Var(2)$ was often taken to be 2. V was often taken to be discrete leading to calculations such as

$P(V > 5) = 1 - P(V \leqslant 5)$.

25b. *[6 marks]*

Markscheme

$\text{E}(V) = \text{E}(3S - 4R)$

$= 3\text{E}(S) - 4\text{E}(R)$ ***M1***

$= 24 - 20 = 4$ ***A1***

$\text{Var}(3S - 4R) = 9\text{Var}(S) + 16\text{Var}(R)$, since R and S are independent random variables ***M1***

$=18 + 16 = 34$ ***A1***

$V \sim \text{N}(4,\ 34)$

$\text{P}(V > 5) = 0.432$ ***A2*** ***N0***

[6 marks]

Examiners report

$E(Y)$ was calculated correctly but many could not go further to find

$Var(Y)$ and $E(Y^2)Var(2)$ was often taken to be 2. V was often taken to be discrete leading to calculations such as

$P(V > 5) = 1 - P(V \leqslant 5)$.

26. *[7 marks]*

Markscheme

Let X denote the number of imperfect glasses in the sample ***(M1)***

For recognising binomial or proportion or Poisson ***A1***

(

$X \sim \mathrm{B}(200,\ p)$ where p-value is the probability of a glass being imperfect)

Let

$\mathrm{H}_0 : p\text{-value} = 0.02$ and $\mathrm{H}_1 : p\text{-value} > 0.02$ ***A1A1***

EITHER

p-value = 0.0493 ***A2***

Using the binomial distribution

$p\text{-value} = 0.0493 > 0.01$ we accept H_0 ***R1***

OR

p-value = 0.0511 ***A2***

Using the Poisson approximation to the binomial distribution since

$p\text{-value} = 0.0511 > 0.01$ we accept H_0 ***R1***

OR

p-value = 0.0217 ***A2***

Using the one proportion z-test since

$p\text{-value} = 0.0217 > 0.01$ we accept H_0 ***R1***

Note: Use of critical values is acceptable.

[7 marks]

Examiners report

Many candidates used a t-test on this question. This was possibly because the sample was large enough to approximate normality of a proportion. The need to use a one-tailed test was often missed. When using the z-test of proportions $p = 0.04$ was often used instead of $p = 0.02$. Not many candidates used the binomial distribution.

27. *[11 marks]*

Markscheme

(a)

$H_0 : d = 0$; $H_1 : d > 0$, where d is the difference in the number of digits remembered ***A1A1***

[2 marks]

(b)

Child	A	B	C	D	E	F	G	H	I	J	K	L
Number of digits remembered on test 1	4	6	4	7	8	5	6	7	6	8	4	7
Number of digits remembered on test 2	7	8	5	5	10	7	7	10	8	6	3	9
Difference (d)	3	2	1	–2	2	2	1	3	2	–2	–1	2

A2

Notes: Award ***A2*** for the correct d values.

Award ***A1*** for one error, ***A0*** for two or more errors.

Use the t-test because the variance is not known ***M1R1***

By GDC

$t = 2.106\ldots$ ***(A2)***

EITHER

p-value = 0.0295 (accept any value that rounds to this number) ***A2***

Since $0.0295 < 0.05$ there is evidence that practice sessions improve ability to memorize digits ***R1***

OR

The critical value of t is 1.796 ***A2***

Since $2.106... > 1.796$ there is evidence that practice sessions improve ability to memorize digits ***R1***

Note: Award ***M1R1A1A1R1*** for testing equality of means ($t = -1.46$, p-value = 0.08) .

[9 marks]

Total [11 marks]

Examiners report

Although this question was reasonably well done the hypotheses were often not stated precisely and the fact that the two data sets were dependent escaped many candidates.

28. *[16 marks]*

Markscheme

(a) With

$H_0,\ \bar{X} \sim N\left(13, \frac{3}{2}\right) = N(13,\ 1.5)$ ***(M1)(A1)***

(i) 5 % for N(0,1) is 1.645

so

$\frac{\bar{x}-13}{\sqrt{1.5}} = 1.645$ ***(M1)(A1)***

$\bar{x} = 13 + 1.645\sqrt{1.5}$

$= 15.0\ \ (3 \text{ s.f.})$ ***A1 N0***

$[15.0,\ \infty[$

(ii) 1% for N(0, 1) is 2.326

so

$\frac{\bar{x}-13}{\sqrt{1.5}} = 2.326$ ***(M1)(A1)***

$\bar{x} = 13 + 2.326\sqrt{1.5}$

$= 15.8\ \ (3 \text{ s.f., accept } 15.9)$ ***A1 N0***

$[15.8,\ \infty[$

[8 marks]

(b) (i)

$\beta = P(\bar{X} < 15.0147)$ ***M1***

$= 0.440$ ***A2***

(ii)

$\beta = P(\bar{X} < 15.8488)$ ***M1***

$= 0.702$ ***A2***

[6 marks]

(c) The probability of a Type II error increases when the probability of a Type I error decreases. ***R2***

[2 marks]

Total [16 marks]

Examiners report

This question proved to be the most difficult. The range of solutions ranged from very good to very poor. Many students thought that $P(TypeI) = 1 - P(TypeII)$ when in fact

$1 - P(TypeII)$ is the power of the test.

29.

[9 marks]

Markscheme

(a)

$n = 7$, sample mean $= 35$ ***(A1)***

$s_{n-1}^2 = \frac{\sum (x-35)^2}{6} = 322$ ***(M1)A1***

[3 marks]

(b) null hypothesis

$\mathrm{H}_0 : \mu = 42.3$ ***A1***

alternative hypothesis

$\mathrm{H}_1 : \mu < 42.3$ ***A1***

using one-sided *t*-test

$|t_{\text{calc}}| = \sqrt{7}\frac{42.3-35}{\sqrt{322}} = 1.076$ ***(M1)(A1)***

with 6 degrees of freedom ,

$t_{\text{crit}} = 1.440 > 1.076$

$(\text{or } p\text{-value } = 0.162 > 0.1)$ ***A1***

we conclude that there is no justification for cutting down the cedar trees ***R1*** ***N0***

Note: ***FT*** on their *t* or *p*-value.

[6 marks]

Total [9 marks]

Examiners report

This question was generally well attempted as an example of the *t*-test. Very few used the *Z* statistic, and many found *p*-values.

30. *[12 marks]*

Markscheme

(a) the distribution is NB(3, 0.09) ***(M1)(A1)***

the probability is

$\binom{24}{2} 0.91^{22} \times 0.09^3 = 0.0253$ ***(M1)(A1)A1***

[5 marks]

(b) P(Heating increased on

n^{th} day)

$\binom{n-1}{2} 0.91^{n-3} \times 0.09^3$ ***(M1)(A1)(A1)***

by trial and error $n = 23$ gives the maximum probability ***(M1)A3***

(neighbouring values: 0.02551 ($n = 22$) ; 0.02554 ($n = 23$) ; 0.02545 ($n = 24$))

[7 marks]

Total [12 marks]

Examiners report

Most candidates understood the context of this question, and the negative binomial distribution was usually applied, albeit occasionally with incorrect parameters. Good solutions were seen to part(b), using lists in their GDC or trial and error.

31. *[15 marks]*

Markscheme

(a) $f(x)$is even (symmetrical about the origin) ***(M1)***

$\mathrm{E}(X) = 0$ ***A1***

$\mathrm{Var}(X) = \mathrm{E}(X^2) = \int_{-0.005}^{0.005} 100x^2 \mathrm{d}x$ ***(M1)(A1)***

$= 8.33 \times 10^{-6} \left(\text{accept } 0.83 \times 10^{-5} \text{ or } \frac{1}{120\,000}\right)$ ***A1***

[5 marks]

(b) rounding errors to 2 decimal places are uniformly distributed ***R1***

and lie within the interval

$-0.005 \leqslant x < 0.005.$ ***R1***

this defines X ***AG***

[2 marks]

(c) (i) using the symbol y to denote the error in the sum of 20 real numbers each rounded to 2 decimal places

$-0.1 \leqslant y (= 20 \times x) < 0.1$ ***A1***

(ii)

$Y \approx \mathrm{N}(20 \times 0,\ 20 \times 8.3 \times 10^{-6}) = \mathrm{N}(0,\ 0.00016)$ ***(M1)(A1)***

$\mathrm{P}\left(|Y| > 0.01\right) = 2\left(1 - \mathrm{P}(Y < 0.01)\right)$ ***(M1)(A1)***

$= 2\left(1 - \mathrm{P}\left(Z < \frac{0.01}{0.0129}\right)\right)$

$= 0.44$ to 2 decimal places ***A1*** ***N4***

(iii) it is assumed that the errors in rounding the 20 numbers are independent ***R1***

and, by the central limit theorem, the sum of the errors can be modelled approximately by a normal distribution ***R1***

[8 marks]

Total [15 marks]

Examiners report

This was the only question on the paper with a conceptually ‘hard’ final part. Part(a) was generally well done, either by integration or by use of the standard formulae for a uniform distribution. Many candidates were not able to provide convincing reasoning in parts (b) and (c)(iii). Part(c)(ii), the application of the Central Limit Theorem was only very rarely tackled competently.

32. *[11 marks]*

Markscheme

(a)

$A \sim \mathrm{N}(30,\ 3^2)$

$\mathrm{P}(A > 35) = 0.0478$ ***(M1)A1***

[2 marks]

(b) let

$X = B_1 + B_2 + B_3 + B_4 + B_5$

$\mathrm{E}(X) = 5\mathrm{E}(B) = 60$ ***A1***

$\mathrm{Var}(X) = 5\mathrm{Var}(B) = 20$ ***(M1)A1***

$\mathrm{P}(X < 70) = 0.987$ ***A1***

[4 marks]

(c) let

$Y = A - 2B$ ***(M1)***

$\mathrm{E}(Y) = \mathrm{E}(A) - 2\mathrm{E}(B) = 6$ ***A1***

$\mathrm{Var}(Y) = \mathrm{Var}(A) + 4\mathrm{Var}(B) = 25$ ***(M1)A1***

$\mathrm{P}(Y > 0) = 0.885$ ***A1***

[5 marks]

Total [11 marks]

Examiners report

Most candidates were able to access this question, but weaker candidates did not always realise that parts (b) and (c) were testing different things. Part (b) proved the hardest with a number of candidates not understanding how to find the variance of the sum of variables.

33.

[13 marks]

Markscheme

(a)

$\bar{x} = \frac{\sum x}{n} = 1.2$ ***(A1)***

$s_{n-1}^2 = 0.524\ldots$ ***(A1)***

it is a one tailed test

$H_0 : \mu = 1.1,\ H_1 : \mu > 1.1$ ***A1***

EITHER

$t = \frac{1.2-1.1}{\sqrt{\frac{0.524\ldots}{15}}} = 0.535$ ***(M1) A1***

$v = 14$ ***(A1)***

$t_{crit} = 1.761$ ***A1***

since

$0.535 < t_{crit}$ we accept

H_0 that there is no increase in the amount of the chemical ***R1***

OR

$p = 0.301$ ***A4***

since

$p > 0.05$ we accept

H_0 that there is no increase in the amount of the chemical ***R1***

[8 marks]

(b) 90 % confidence interval

$= 1.18 \pm 1.645\sqrt{\frac{0.0256}{15}}$ ***(M1)A1A1A1***

$= [1.11,\ 1.25]$ ***A1*** ***N5***

[5 marks]

Total [13 marks]

Examiners report

This question also proved accessible to a majority of candidates with many wholly correct or nearly wholly correct answers seen. A few candidates did not recognise that part (a) was a *t*-distribution and part (b) was a Normal distribution, but most recognised the difference. Many candidates received an accuracy penalty on this question for not giving the final answer to part (b) to 3 significant figures.

34. *[14 marks]*

Markscheme

(a) geometric distribution ***A1***

[1 mark]

(b) let R be the event throwing the disc and it landing on red and

let B be the event throwing the disc and it landing on blue

$\mathrm{P}(X=1)=p=\mathrm{P}\left(1B \text{ and } (n-1)R \text{ or } 1R \text{ and } (n-1)B\right)$ ***(M1)***

$=n\times\frac{1}{2}\times\left(\frac{1}{2}\right)^{n-1}+n\times\frac{1}{2}\times\left(\frac{1}{2}\right)^{n-1}$ ***(A1)***

$=\frac{n}{2^{n-1}}$ ***A1***

hence

$\mathrm{P}(X=x)=\frac{n}{2^{n-1}}\left(1-\frac{n}{2^{n-1}}\right)^{x-1},\ (x\geqslant 1)$ ***A1***

Notes:

$x\geqslant 1$ not required for final ***A1***.

Allow ***FT*** for final ***A1***.

[4 marks]

(c)

$\mathrm{E}(X)=\frac{1}{p}$

$=\frac{2^{n-1}}{n}$ ***A1***

[1 mark]

(d) when

$n=7$,

$\mathrm{P}(X=x)=\left(1-\frac{7}{64}\right)^{x-1}\times\frac{7}{64}$ ***(M1)***

$=\frac{7}{64}\times\left(\frac{57}{64}\right)^{x-1}$

$\mathrm{P}(X\leqslant k)=\sum\limits_{x=1}^{k}\frac{7}{64}\times\left(\frac{57}{64}\right)^{x-1}$ ***(M1)(A1)***

$\Rightarrow\frac{7}{64}\times\frac{1-\left(\frac{57}{64}\right)^{k}}{1-\frac{57}{64}}>0.5$ ***(M1)(A1)***

$\Rightarrow 1-\left(\frac{57}{64}\right)^{k}>0.5$

$\Rightarrow\left(\frac{57}{64}\right)^{k}<0.5$

$\Rightarrow k>\frac{\log 0.5}{\log\frac{57}{64}}$ ***(M1)***

$\Rightarrow k>5.98$ ***(A1)***

$\Rightarrow k=6$ ***A1***

Note: Tabular and other GDC methods are acceptable.

[8 marks]

Total [14 marks]

Examiners report

This question was found difficult by the majority of candidates and few fully correct answers were seen. Few candidates were able to find

$\mathrm{P}(X=x)$ in terms of n and x and many did not realise that the last part of the question required them to find the sum of a series. However, better candidates received over 75% of the marks because the answers could be followed through.

35a.

[2 marks]

Markscheme

Interval	Frequency
$]1.0, 1.1]$	6
$]1.1, 1.2]$	28
$]1.2, 1.3]$	18
$]1.3, 1.4]$	14
$]1.4, 1.5]$	10
$]1.5, 1.6]$	4

A2

[2 marks]

Examiners report

[N/A]

35b.

[2 marks]

Markscheme

$\mu = 1.26,\ \sigma = 0.133$ ***A1A1***

[2 marks]

Examiners report

[N/A]

35c.

[2 marks]

Markscheme

no because the normal distribution is symmetric and these data are not ***R2***

[2 marks]

Examiners report

[N/A]

36a.

[3 marks]

Markscheme

unbiased estimate of

$\mu = 122$ ***A1***

unbiased estimate of

$\sigma^2 = 4.4406\ldots^2 = 19.7$ ***(M1)A1***

Note: Award ***(M1)A0*** for 4.44.

[3 marks]

Examiners report

[N/A]

36b. *[2 marks]*

Markscheme

the 99 % confidence interval for

μ is [118, 126] ***A1A1***

[2 marks]

Examiners report

[N/A]

36c. *[5 marks]*

Markscheme

(i)

$\mathrm{H}_0 : \mu = 125$; $\mathrm{H}_1 : \mu < 125$ ***A1***

(ii) p-value = 0.0220 ***A2***

(iii) the shopper's claim is supported because

$0.0220 < 0.05$ ***A1R1***

[5 marks]

Examiners report

[N/A]

37a. *[2 marks]*

Markscheme

$\mathrm{P}(X = 1) = \frac{1}{3}$ ***A1***

$\mathrm{P}(X = 2) = \frac{2}{3} \times \frac{1}{4}$ ***A1***

$= \frac{1}{6}$ ***AG***

[2 marks]

Examiners report

[N/A]

37b. *[6 marks]*

Markscheme

$G(t) = \frac{1}{3}t + \frac{2}{3} \times \frac{1}{4}t^2 + \frac{2}{3} \times \frac{3}{4} \times \frac{1}{3}t^3 + \frac{2}{3} \times \frac{3}{4} \times \frac{2}{3} \times \frac{1}{4}t^4 + \ldots$ ***M1A1***

$= \frac{1}{3}t\left(1 + \frac{1}{2}t^2 + \ldots\right) + \frac{1}{6}t^2\left(1 + \frac{1}{2}t^2 + \ldots\right)$ ***M1A1***

$= \frac{\frac{t}{3}}{1 - \frac{t^2}{2}} + \frac{\frac{t^2}{6}}{1 - \frac{t^2}{2}}$ ***A1A1***

$= \frac{2t + t^2}{6 - 3t^2}$ ***AG***

[6 marks]

Examiners report

[N/A]

37c. *[4 marks]*

Markscheme

$G'(t) = \frac{(2+2t)(6-3t^2)+6t(2t+t^2)}{(6-3t^2)^2}$ ***M1A1***

$\mathrm{E}(X) = G'(1) = \frac{10}{3}$ ***M1A1***

[4 marks]

Examiners report

[N/A]

38a. *[5 marks]*

Markscheme

we are given that

$M \sim \mathrm{N}(30, 9)$ and

$F \sim \mathrm{N}(20, 6.25)$

let

$X = M - 2F$; $X \sim \mathrm{N}($

$-10,$

$34)$ ***M1A1A1***

we require

$\mathrm{P}(X > 0)$ ***(M1)***

= 0.0432 ***A1***

[5 marks]

Examiners report

[N/A]

38b. *[4 marks]*

Markscheme

let

$Y = M_1 + M_2 + F_1 + F_2 + F_3 + F_4 + F_5$; $Y \sim \mathrm{N}(160, 49.25)$ ***M1A1A1***

we require

$\mathrm{P}(Y < 175) = 0.984$ ***A1***

[4 marks]

Examiners report

[N/A]

39a. *[1 mark]*

Markscheme

$\mathrm{H}_0 : \rho = 0$; $\mathrm{H}_1 : \rho > 0$ ***A1***

[1 mark]

Examiners report

[N/A]

39b. *[5 marks]*

Markscheme

(i) correlation coefficient = 0.905 ***A2***

p-value

$= 2.61 \times 10^{-5}$ ***A2***

(ii) very strong evidence to indicate a positive association between marks in Mechanics and marks in Statistics ***R1***

[5 marks]

Examiners report

[N/A]

39c. *[4 marks]*

Markscheme

the regression line of y on x is

$y = 8.71 + 0.789x$ ***(M1)A1***

George's estimated mark on Paper 2

$= 8.71 + 0.789 \times 63$ ***(M1)***

= 58 ***A1***

[4 marks]

Examiners report

[N/A]

39d. *[5 marks]*

Markscheme

$t = r\sqrt{\frac{n-2}{1-r^2}} = 2.3019\ldots$ ***M1A1***

degrees of freedom = 14 ***(A1)***

p-value

$= 0.0186\ldots$ ***A1***

at the 1 % significance level, this does not indicate a positive association between the marks in Physics and Chemistry ***R1***

[5 marks]

Examiners report

[N/A]