| | | | |
|---|---|---|---|
| **(i)** | $P(X = 1) = 8 \times 0.1^1 \times 0.9^7$<br>$= 0.383$ | M1 for binomial probability $P(X=1)$<br>A1 (at least 2sf) CAO | **2** |
| **(ii)** | $\lambda = 30 \times 0.1 = 3$<br>(*A*) $P(X = 6) = e^{-3}\frac{3^6}{6!} = 0.0504$(3 s.f.)<br>or from tables $= 0.9665 - 0.9161 = 0.0504$<br>(*B*) Using tables: $P(X \geq 8) = 1 - P(X \leq 7)$<br>$= 1 - 0.9881 = 0.0119$ | B1 for mean SOI<br>M1 for calculation or use of tables to obtain $P(X=6)$<br>A1 (at least 2sf) CAO<br>M1 for correct probability calc'<br>A1 (at least 2sf) CAO | **1**<br>**2**<br>**2** |
| **(iii)** | *n* is large and *p* is small | B1, B1<br>Allow appropriate numerical ranges | **2** |
| **(iv)** | $\mu = np = 120 \times 0.1 = 12$<br>$\sigma^2 = npq = 120 \times 0.1 \times 0.9 = 10.8$ | B1<br>B1 | **2** |
| **(v)** | $P(X > 15.5) = P\left(Z > \frac{15.5-12}{\sqrt{10.8}}\right)$<br>$= P(Z > 1.065) = 1 - \Phi(1.065) = 1 - 0.8566$<br>$= 0.1434$<br>NB Allow full marks for use of N(12,12) as an approximation to Poisson(12) leading to $1 - \Phi(1.010) = 1 - 0.8438 = 0.1562$ | B1 for correct continuity correction.<br>M1 for probability using correct tail<br>A1 **cao, (**but FT wrong or omitted CC) | **3** |
| **(vi)** | From tables $\Phi^{-1}(0.99) = 2.326$<br>$\frac{x + 0.5 - 12}{\sqrt{10.8}} \geq 2.326$<br>$x = 11.5 + 2.326 \times \sqrt{10.8} \geq 19.14$<br>So 20 breakfasts should be carried<br>NB Allow full marks for use of N(12,12) leading to<br>$x \geq 11.5 + 2.326 \times \sqrt{12} = 19.56$ | B1 for 2.326 seen<br>M1 for equation in *x* and positive *z*-value<br>A1 CAO (condone 19.64)<br>A1FT for rounding appropriately (i.e. round up if c.c. used o/w rounding should be to nearest integer) | **4** |
| | | | **18** |

## Question 2

| | | | |
|---|---|---|---|
| **(i)** | $X \sim N(49.7, 1.6^2)$<br>(*A*) $P(X > 51.5) = P\left(Z > \frac{51.5 - 49.7}{1.6}\right)$<br>$= P(Z > 1.125)$<br>$= 1 - \Phi(1.125) = 1 - 0.8696 = 0.1304$<br>(*B*) $P(X < 48.0) = P\left(Z < \frac{48.0 - 49.7}{1.6}\right)$<br>$= P(Z < -1.0625) = 1 - \Phi(1.0625)$<br>$= 1 - 0.8560 = 0.1440$<br>$P(48.0 < X < 51.5) = 1 - 0.1304 - 0.1440 = 0.7256$ | M1 for standardizing<br>M1 for prob. calc.<br>A1 (at least 2 s.f.)<br>M1 for appropriate prob' calc.<br>A1 (0.725 – 0.726) | **5** |
| **(ii)** | P(one over 51.5, three between 48.0 and 51.5)<br>$= \binom{4}{1} \times 0.7256 \times 0.2744^3 = 0.0600$ | M1 for coefficient<br>M1 for $0.7256 \times 0.2744^3$<br>A1 FT (at least 2 sf) | **3** |
| **(iii)** | From tables,<br>$\Phi^{-1}(0.60) = 0.2533$, $\Phi^{-1}(0.30) = -0.5244$<br>$49.0 = \mu + 0.2533\,\sigma$<br>$47.5 = \mu - 0.5244\,\sigma$<br>$1.5 = 0.7777\,\sigma$<br>$\sigma = 1.929$, $\mu = 48.51$ | B1 for 0.2533 or 0.5244 seen<br>M1 for at least one correct equation $\mu$ & $\sigma$<br>M1 for attempt to solve two correct equations<br>A1 CAO for both | **4** |
| **(iv)** | Where $\mu$ denotes the mean circumference of the entire population of organically fed 3-year-old boys.<br>$n = 10$,<br>Test statistic $Z = \frac{50.45 - 49.7}{1.6/\sqrt{10}} = \frac{0.75}{0.5060} = 1.482$<br>10% level 1 tailed critical value of $z$ is 1.282<br>1.482 > 1.282 so significant.<br>There is sufficient evidence to reject $H_0$ and conclude that organically fed 3-year-old boys have a higher mean head circumference. | E1<br>M1<br>A1(at least 3sf)<br>B1 for 1.282<br>M1 for comparison leading to a conclusion<br>A1 for conclusion in context | **6** |
| | | | **18** |

## Question 3

| | | | |
|---|---|---|---|
| **(i)** | **EITHER:**<br>$S_{xy} = \Sigma xy - \frac{1}{n}\Sigma x \Sigma y = 6235575 - \frac{1}{10} \times 4715 \times 13175$<br>$= 23562.5$<br>$S_{xx} = \Sigma x^2 - \frac{1}{n}(\Sigma x)^2 = 2237725 - \frac{1}{10} \times 4715^2 =$<br>$14602.5$<br>$S_{yy} = \Sigma y^2 - \frac{1}{n}(\Sigma y)^2 = 17455825 - \frac{1}{10} \times 13175^2 =$<br>$97762.5$<br>$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = \frac{23562.5}{\sqrt{14602.5 \times 97762.5}} = 0.624$<br>**OR:**<br>cov $(x,y) = \frac{\sum xy}{n} - \overline{x}\,\overline{y} = 6235575/10 - 471.5 \times 1317.5$<br>$= 2356.25$<br>rmsd$(x) = \sqrt{\frac{S_{xx}}{n}} = \sqrt{(14602.5/10)} = \sqrt{1460.25} = 38.21$<br>rmsd$(y) = \sqrt{\frac{S_{yy}}{n}} = \sqrt{(97762.5/10)} = \sqrt{9776.25} = 98.87$<br>$r = \frac{\text{cov}(x,y)}{rmsd(x)rmsd(y)} = \frac{2356.25}{38.21 \times 98.87} = 0.624$ | M1 for method for $S_{xy}$<br>M1 for method for at least one of $S_{xx}$ or $S_{yy}$<br>A1 for at least one of $S_{xy}$, $S_{xx}$ or $S_{yy}$ correct<br>M1 for structure of $r$<br>A1 (0.62 to 0.63)<br>M1 for method for cov $(x,y)$<br>M1 for method for at least one msd<br>A1 for at least one of $S_{xy}$, $S_{xx}$ or $S_{yy}$ correct<br>M1 for structure of $r$<br>A1 (0.62 to 0.63) | **5** |
| **(ii)** | $H_0$: $\rho = 0$<br>$H_1$: $\rho \neq 0$ (two-tailed test)<br>where $\rho$ is the population correlation coefficient<br>For $n = 10$, 5% critical value = 0.6319<br>Since 0.624 < 0.6319 we cannot reject $H_0$:<br>There is not sufficient evidence at the 5% level to suggest that there is any correlation between length and circumference. | B1 for $H_0$, $H_1$ in symbols<br>B1 for defining $\rho$<br>B1FT for critical value<br>M1 for sensible comparison leading to a conclusion<br>A1 FT for result<br>B1 FT for conclusion in context | **6** |
| **(iii)** | *(A)* This is the probability of rejecting $H_0$ when it is in fact true.<br>*(B)* Advantage of 1% level – less likely to reject $H_0$ when it is true.<br>Disadvantage of 1% level – less likely to accept $H_1$ when $H_0$ is false. | B1 for 'P(reject $H_0$)'<br>B1 for 'when true'<br>B1, B1 Accept answers in context | **2**<br>**2** |

| | | | |
|---|---|---|---|
| | | | |
| **(iv)** | The student's approach is not valid.<br>If a statistical procedure is repeated with a new sample, we should not simply ignore one of the two outcomes.<br>The student could combine the two sets of data into a single set of twenty measurements. | E1<br><br>E1 – allow suitable alternatives.<br>E1 for combining samples. | **3** |
| | | | **18** |

## Question 4

**(i)**

$H_0$: no association between musical preference and age;
$H_1$: some association between musical preference and age;

| **Observed** | | Musical preference | | | **Row totals** |
|---|---|---|---|---|---|
| | | Pop | Classical | Jazz | |
| Age group | Under 25 | 57 | 15 | 12 | **84** |
| | 25 – 50 | 43 | 21 | 21 | **85** |
| | Over 50 | 22 | 32 | 27 | **81** |
| **Column totals** | | **122** | **68** | **60** | **250** |

| **Expected** | | Musical preference | | | **Row totals** |
|---|---|---|---|---|---|
| | | Pop | Classical | Jazz | |
| Age group | Under 25 | 40.992 | 22.848 | 20.160 | **84** |
| | 25 – 50 | 41.480 | 23.120 | 20.400 | **85** |
| | Over 50 | 39.528 | 22.032 | 19.440 | **81** |
| **Column totals** | | **122** | **68** | **60** | **250** |

| **Contributions** | | Musical preference | | | |
|---|---|---|---|---|---|
| | | Pop | Classical | Jazz | |
| Age group | Under 25 | 6.25 | 2.70 | 3.30 | |
| | 25 – 50 | 0.06 | 0.19 | 0.02 | |
| | Over 50 | 7.77 | 4.51 | 2.94 | |
| | | | | | |

$X^2 = 27.74$

Refer to $\chi_4^2$
Critical value at 5% level = 9.488
Result is significant
There is some association between age group and musical preference.
NB if $H_0$ $H_1$ reversed, or 'correlation' mentioned, do not award first B1or final E1

B1

M1 A2 for expected values (at least 1 dp) (allow A1 for at least one row or column correct)

M1 for valid attempt at $(O-E)^2/E$
A1 for all correct

M1dep for summation
A1 for $X^2$ (27.7 – 27.8)

B1 for 4 deg of f

B1 CAO for cv
B1FT
E1 (conclusion in context)

**1**

**7**

**4**

| | | | |
|---|---|---|---|
| **(ii)** | The values of 6.25 and 7.77 show that under 25's have a strong positive association with pop whereas over 50's have a strong negative association with pop.<br>The values of 4.51 and 2.94 show that over 50's have a reasonably strong positive association with both classical and jazz.<br>The values of 2.70 and 3.30 show that under 25's have a reasonably strong negative associations with both classical and jazz.<br>The 25-50 group's preferences differ very little from the overall preferences. | B1, B1<br>for specific reference to a value from the table of contributions followed by an appropriate comment<br>B1, B1 (as above for second value )<br>B1, B1 (as above for third value) | **6** |
| | | | **18** |