A class of students had a sudoku competition. The time taken for each student to complete the sudoku was recorded to the nearest minute and the results are summarised in the table below.

|  |  |  |
| --- | --- | --- |
| Time | Mid-point, *x* | Frequency, f |
| 2 – 8 | 5 | 2 |
| 9 – 12 |  | 7 |
| 13 – 15 | 14 | 5 |
| 16 – 18 | 17 | 8 |
| 19 – 22 | 20.5 | 4 |
| 23 – 30 | 26.5 | 4 |

(You may use ∑f*x*2 = 8603.75)

(*a*) Write down the mid-point for the 9 – 12 interval.

**(1)**

(*b*) Use linear interpolation to estimate the median time taken by the students.

**(2)**

(*c*) Estimate the mean and standard deviation of the times taken by the students.

**(5)**

The teacher suggested that a normal distribution could be used to model the times taken by the students to complete the sudoku.

(*d*) Give a reason to support the use of a normal distribution in this case.

**(1)**

On another occasion the teacher calculated the quartiles for the times taken by the students to complete a different sudoku and found

*Q*1 = 8.5 *Q*2 =13.0 *Q*3 = 21.0

(*e*) Describe, giving a reason, the skewness of the times on this occasion.

**(2)**

**Total 11 marks**

**S1 May 2011 qu.5**

A survey of 100 households gave the following results for weekly income £*y*.

|  |  |  |
| --- | --- | --- |
| Income *y* (£) | Mid-point | Frequency *f* |
| 0 ≤ *y* < 200 | 100 | 12 |
| 200 ≤ *y* < 240 | 220 | 28 |
| 240 ≤ *y* < 320 | 280 | 22 |
| 320 ≤ *y* < 400 | 360 | 18 |
| 400 ≤ *y* < 600 | 500 | 12 |
| 600 ≤ *y* < 800 | 700 | 8 |

(You may use ∑ *fy*2 = 12 452 800)

A histogram was drawn and the class 200 ≤ *y*< 240 was represented by a rectangle of width 2 cm and height 7 cm.

(*a*) Calculate the width and the height of the rectangle representing the class 320 ≤ *y* < 400.

**(3)**

(*b*) Use linear interpolation to estimate the median weekly income to the nearest pound.

**(2)**

(*c*) Estimate the mean and the standard deviation of the weekly income for these data.

**(4)**

One measure of skewness is .

(*d*) Use this measure to calculate the skewness for these data and describe its value.

**(2)**

Katie suggests using the random variable *X* which has a normal distribution with mean 320 and standard deviation 150 to model the weekly income for these data.

(*e*) Find P(240 < *X* < 400).

**(2)**

(*f*) With reference to your calculations in parts (*d*) and (*e*) and the data in the table, comment on Katie’s suggestion.

**(2)**

**Total 15 marks**

**S1 January 2013 qu.5**

A midwife records the weights, in kg, of a sample of 50 babies born at a hospital. Her results are given in the table below.

|  |  |  |
| --- | --- | --- |
| **Weight (*w* kg)** | **Frequency (*f*)** | **Weight midpoint (*x*)** |
| 0 ≤ *w <* 2 | 1 | 1 |
| 2 ≤ *w* < 3 | 8 | 2.5 |
| 3 ≤ *w* < 3.5 | 17 | 3.25 |
| 3.5 ≤ *w* < 4 | 17 | 3.75 |
| 4 ≤ *w* < 5 | 7 | 4.5 |

[You may use  = 611.375]

A histogram has been drawn to represent these data.

The bar representing the weight 2 ≤ *w* < 3 has a width of 1 cm and a height of 4 cm.

(*a*)Calculate the width and height of the bar representing a weight of 3 ≤ *w* < 3.5.

**(3)**

(*b*)Use linear interpolation to estimate the median weight of these babies.

**(2)**

(*c*)(i) Show that an estimate of the mean weight of these babies is 3.43 kg.

(ii) Find an estimate of the standard deviation of the weights of these babies.

**(3)**

Shyam decides to model the weights of babies born at the hospital, by the random variable *W*, where *W* ~ N(3.43, 0.652).

(*d*)Find P(*W* < 3).

**(3)**

(*e*)With reference to your answers to (*b*), (*c*)(i) and (*d*)comment on Shyam’s decision.

**(3)**

A newborn baby weighing 3.43 kg is born at the hospital.

(*f*)Without carrying out any further calculations, state, giving a reason, what effect the addition of this newborn baby to the sample would have on your estimate of the

(i) mean,

(ii) standard deviation.

**(3)**

**Total 17 marks**

**S1 June 2016 qu.5**

**Mark scheme**

|  |  |  |
| --- | --- | --- |
| Question Number | Scheme | Marks |
| **5. (a)** | 10.5 | B1 |
|  |  | (1) |
| **(b)** |  | M1 |
|  | = 15.875 or 16.0625 | A1 |
|  |  | (2) |
| **(c)** | = 15.9 () [ Accept ] | M1, A1 |
|  | ,= 5.78 (accept *s* = 5.88) | M1A1ft, A1 |
|  |  | (5) |
| **(d)** | Since mean and median are similar (or equal or very close) a normal distribution may be suitable. [Allow mean or median close to mode/modal class] | B1 |
|  |  | (1) |
| **(e)\*** |  | M1 |
|  | Therefore positive skew | A1 |
|  |  | (2) |
|  |  | **(11 marks)** |

Parts (d) and (e) assess spec point 6 Normal distribution.

\*Part (e) is AO3

| Question Number | | Scheme | Marks |
| --- | --- | --- | --- |
| **5.** (a) | Width = 4 (cm) | | B1 |
|  | Area of 14 cmrepresents frequency 28 and area of 4*h* represents 18 | | M1 |
|  | Or  (o.e.) ***h* = 2.25** (cm) | | A1 |
|  |  | | (3) |
| (b) |  | | M1 |
|  | = 276.36... () ( (£)**276**  < *m* < (£)276.5) | | A1 |
|  |  | | (2) |
| (c) | leading to | | M1A1 |
|  | = 157 .07... (awrt **157**) Allow *s* = 157.86... | | M1A1 |
|  |  | | (4) |
| (d) | Skewness = 0.764... (awrt **0.76** or **0.75**) | | B1 |
|  | [If *n*+1 used in (b) and *m* = £278 accept awrt 0.73 or 0.72] | |  |
|  | Positive skew | | B1ft |
|  |  | | (2) |
| (e) |  | | M1 |
|  | P(240< *X* < 400) = **0.40 ~ 0.41** | | A1 |
|  |  | | (2) |
| (f)\* | (e) suggests a reasonable fit for this range BUT | | B2/1/0 |
|  | (d) since skew it will not be a good fit overall | | (2) |
|  |  | | **15** |

Parts (e) and (f) assess spec point 6 Normal distribution.

\*Part (f) is AO3

|  |  |  |
| --- | --- | --- |
| **Question** | **Scheme** | **Marks** |
| **5. (a)** | Width = **0.5** (cm) | B1 |
|  | e.g. 4 [cm2] represents 8 babies or frequency densities are 8 and 34 | M1 |
|  | Height = **17** (cm) | A1 |
|  |  | **(3)** |
| **(b)** | , or  = awrt **3.47** (allow ) | M1 A1 |
|  |  | **(2)** |
| **(c)(i)** | ,  (\*) | B1 cso |
| **(ii)** | = 0.680147… = awrt **0.680** (accept 0.68) | M1 A1 |
|  |  | **(3)** |
| **(d)** |  | M1 |
|  | = 1 – 0.7454 (tables) | M1 |
|  | = 0.2546 awrt **0.2540.255** | A1 |
|  |  | **(3)** |
| **(e)** | (b) and (c)(i) mean ≠ med or skew or meanmedian or no skew and comment | B1 |
|  | (d) = 0.254 or 0.255 compare data = 0.18 (or 12.7 compared with 9) | B1 |
|  | 0.18 different from 0.25 so normal not good  or 0.18 similar to -.25 so normal is ok | dB1 |
|  |  | **(1)** |
| **(f)\*(i)** | No change in mean (since weight is the same) | B1 |
| **(ii)** | s.d. will decrease (Extra value is at “centre” so data more concentrated) | B1 |
|  | Both statements correct and correct reasons for each | dB1 **(3)** |
|  |  | **[17]** |

Parts (d) and (e) assess spec point 6 Normal distribution.

\*Part (e) is AO3