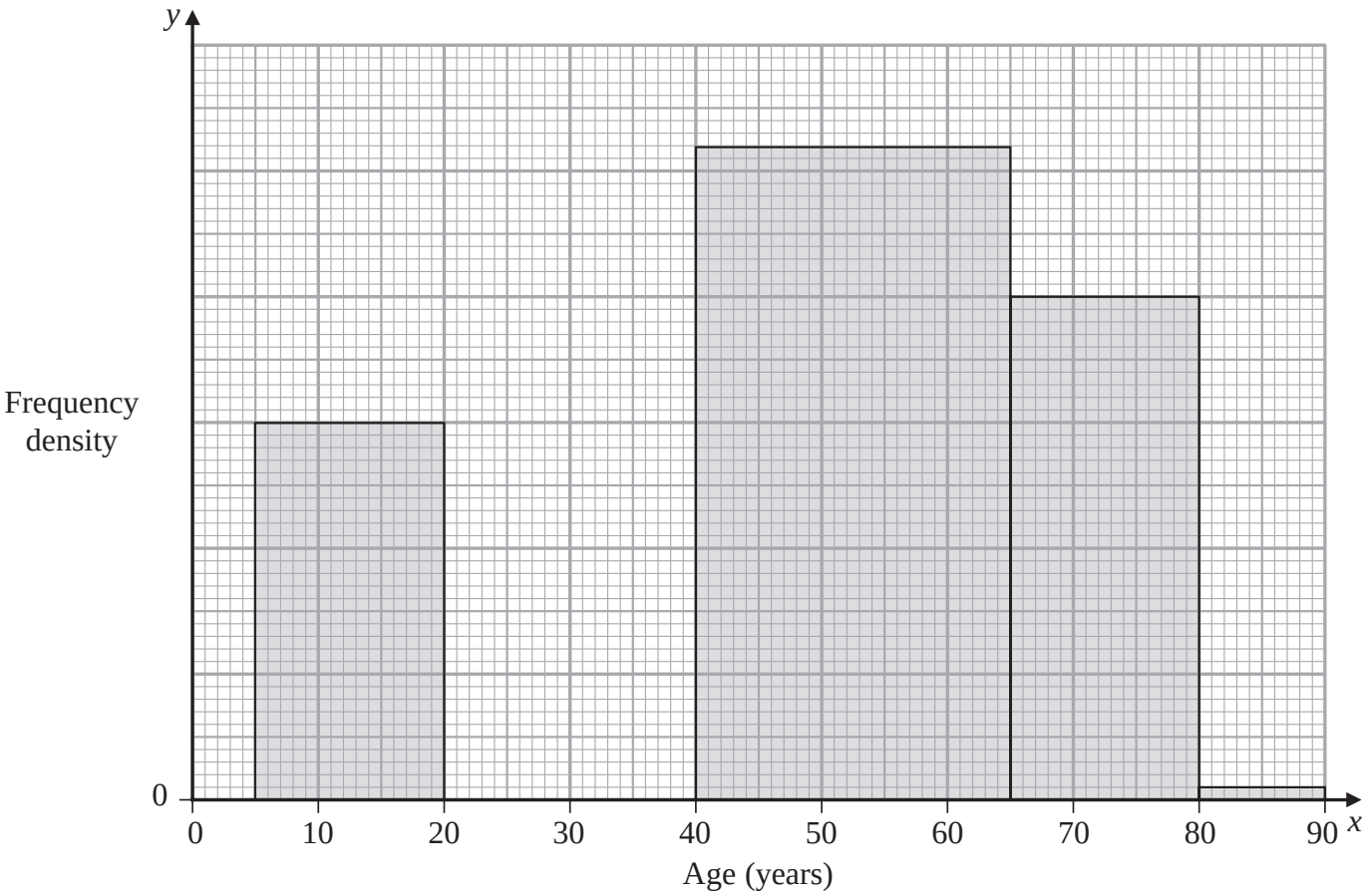
**1.** The partially completed table and partially completed histogram give information about the ages of passengers on an airline.

There were no passengers aged 90 or over.

| **Age ($x$ years)** | $0 \leqslant x < 5$ | $5 \leqslant x < 20$ | $20 \leqslant x < 40$ | $40 \leqslant x < 65$ | $65 \leqslant x < 80$ | $80 \leqslant x < 90$ |
|---|---|---|---|---|---|---|
| **Frequency** | 5 | 45 | 90 | | | 1 |




(a) Complete the histogram.

**(3)**

(b) Use linear interpolation to estimate the median age.

**(4)**

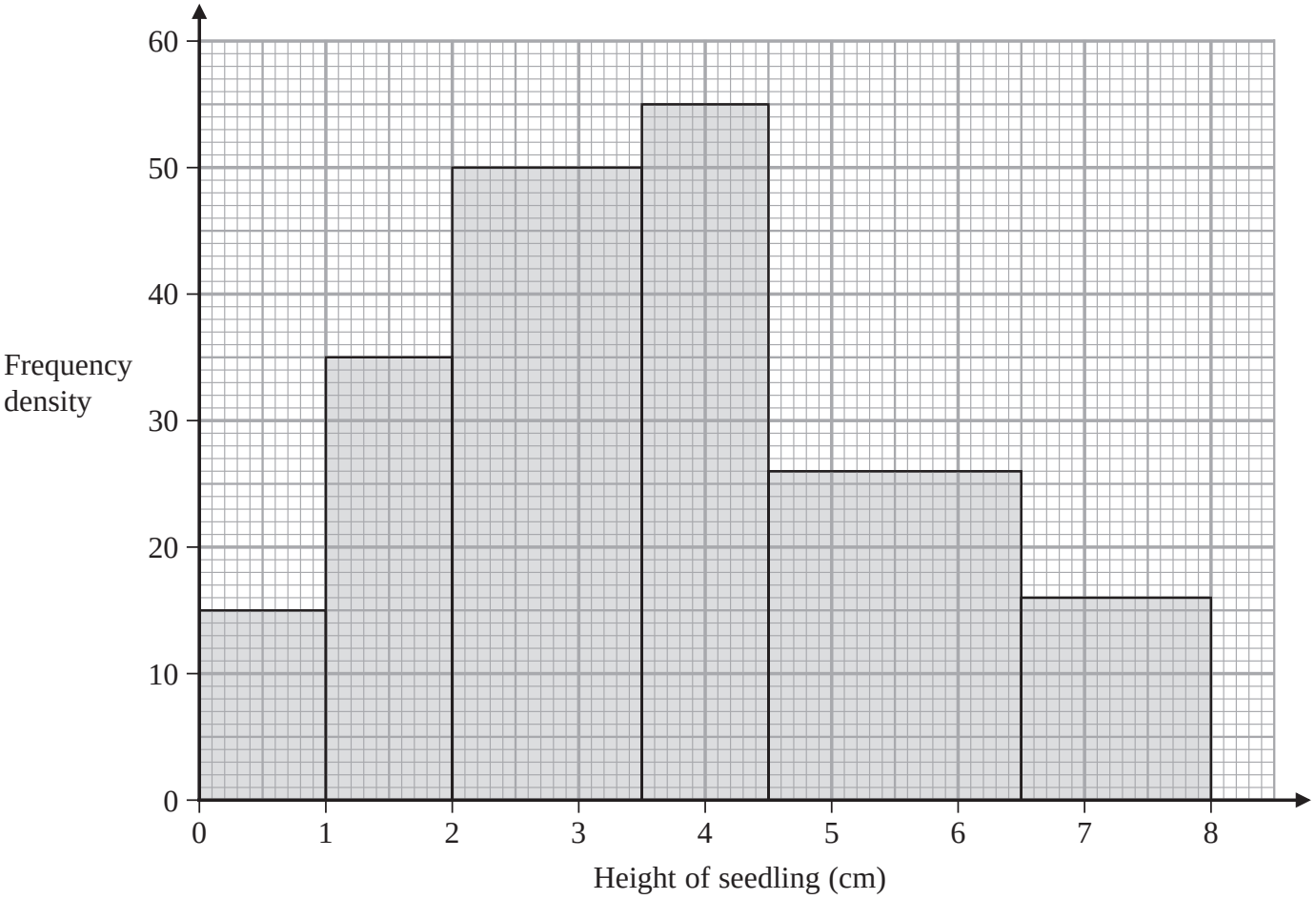
An outlier is defined as a value greater than $Q_3 + 1.5 \times$ interquartile range.

Given that $Q_1 = 27.3$ and $Q_3 = 58.9$

(c) determine, giving a reason, whether or not the oldest passenger could be considered as an outlier.

**(2)**

**2.** The histogram summarises the heights of 256 seedlings two weeks after they were planted.




(a) Use linear interpolation to estimate the median height of the seedlings.

**(4)**

Chris decides to model the **frequency density** for these 256 seedlings by a curve with equation

$$y = kx(8 - x) \qquad 0 \leqslant x \leqslant 8$$

where $k$ is a constant.

(b) Find the value of $k$

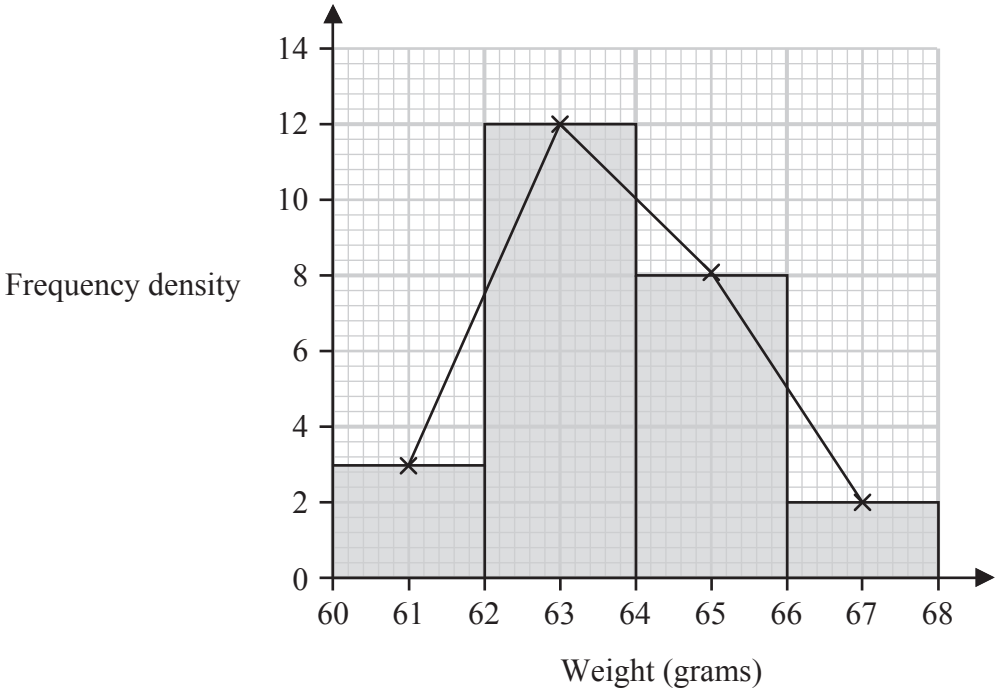
**(3)**

Using this model,

(c) write down the median height of the seedlings.

**(1)**

**3.** The histogram and its frequency polygon below give information about the weights, in grams, of 50 plums.




(a) Show that an estimate for the mean weight of the 50 plums is 63.72 grams.

**(2)**

(b) Calculate an estimate for the standard deviation of the 50 plums.

**(2)**

Later it was discovered that the scales used to weigh the plums were broken.

Each plum actually weighs 5 grams less than originally thought.

(c) State the effect this will have on the estimate of the standard deviation in part (b). Give a reason for your answer.

**(1)**

**4.** A medical researcher is studying the number of hours, $T$, a patient stays in hospital following a particular operation.

The histogram on the page opposite summarises the results for a random sample of 90 patients.

(a) Use the histogram to estimate $P(10 < T < 30)$

**(2)**

For these 90 patients the time spent in hospital following the operation had

- a mean of 14.9 hours
- a standard deviation of 9.3 hours

Tomas suggests that $T$ can be modelled by $N(14.9, 9.3^2)$

(b) With reference to the histogram, state, giving a reason, whether or not Tomas' model could be suitable.

**(1)**

Xiang suggests that the frequency polygon based on this histogram could be modelled by a curve with equation

$$y = kxe^{-x} \quad 0 \leqslant x \leqslant 4$$

where

- $x$ is measured in **tens of hours**
- $k$ is a constant

(c) Use algebraic integration to show that

$$\int_0^n xe^{-x}\,dx = 1 - (n+1)e^{-n}$$

**(4)**

(d) Show that, for Xiang's model, $k = 99$ to the nearest integer.

**(3)**

(e) Estimate $P(10 < T < 30)$ using

(i) Tomas' model of $T \sim N(14.9, 9.3^2)$

**(1)**

(ii) Xiang's curve with equation $y = 99xe^{-x}$ and the answer to part (c)

**(2)**

The researcher decides to use Xiang's curve to model $P(a < T < b)$

(f) State one limitation of Xiang's model.

**(1)**

**Question 4 continued**