Exercise 6A

1 a A census observes or measures every member of a population.

b An advantage is it that it will give a completely accurate result.

A disadvantage could be any one from:
It would be time consuming.
It would be expensive.
It would be difficult to process the data.

2 a The testing process will destroy the harness, so a census would destroy *all* the harnesses, meaning that there would be no harnesses left to sell.

b The claim is misleading. The figure of 250 kg is the mean and median load at which the harnesses in the sample break. So we would expect half of the harnesses to break at a load of less than 250 kg.

c Test a larger number of harnesses.

3 a Any one from:
It would be time consuming.
It would be expensive.
It would be difficult to process the data.

b A list of residents.

c Each individual resident.

4 a The testing process would destroy the switches, so a census would destroy *all* the switches, meaning that there would be no switches left to sell.

b The mean is 19 615.4, less than the stated average. One of the switches failed after significantly fewer operations, which suggests that the median of 22 921 might be a better average to take, as it is not affected by outliers. The data therefore supports the company's claim.

c Test a larger number of switches.

5 a All the mechanics in the garage.

b Everyone’s views will be known.

Exercise 6B

1 The mean is taken from the sample so it is a statistic.

2 **i** and **ii** are statistics

3 **a** All the hairdressers who work for the chain of hairdressing shops.
The proportion p of the staff who are happy to wear an apron.

b This is a binomial distribution, since we are only interested in two options – whether the hairdressers are happy or not.

4 **a** Po(3)

b $P(X < 2) = \frac{e^{-3}3^0}{0!} + \frac{e^{-3}3^1}{1!}$
$= 0.199$ (3 s.f.)

5 **a** $\mu = E(X) = 0.5 \times 50 + 0.25 \times 20 + 0.25 \times 10$
$= 32.5$
$\text{Var}(X) = 0.5 \times 50^2 + 0.25 \times 20^2 + 0.25 \times 10^2 - 32.5^2$
$= 318.75$

b (50, 50), (50, 20), (20, 50), (50, 10), (10, 50), (20, 20), (20, 10), (10, 20), (10, 10)

c

x	50	35	30	20	15	10
$P(X = x)$	0.25	0.25	0.25	0.0625	0.125	0.0625

6 **a** $\mu = E(X) = 0.4 \times 16 + 0.5 \times 20 + 0.1 \times 30$
$= 19.4$
$\text{Var}(X) = 0.4 \times 16^2 + 0.5 \times 20^2 + 0.1 \times 30^2 - 19.4^2$
$= 16.04$

b (16, 16), (16, 20), (20, 16), (16, 30), (30, 16), (30, 30), (30, 20) (20, 30) (20, 20)

c

x	16	18	20	23	25	30
$P(X = x)$	0.16	0.4	0.25	0.08	0.1	0.01

7 a $\mu = \text{E}(X) = 3 \times 0.6 + 2 \times 0.4$
$= 2.6$
$\text{Var}(X) = 3^2 \times 0.6 + 2^2 \times 0.4 - 2.6^2$
$= 0.24$

b (3, 3, 3), (3, 3, 2), (3, 2, 3), (2, 3, 3), (3, 2, 2), (2, 3, 2), (2, 2, 3), (2, 2, 2)

c The sampling distribution for $\overline{X}$ is shown in the table.

x	3	$\frac{8}{3}$	$\frac{7}{3}$	2
$\text{P}(X = x)$	0.216	0.432	0.288	0.064

d The mode can only take the values 2 and 3.
Let $\text{P}(2) = p = 0.4$
Let $\text{P}(3) = q = 0.6$
$\text{P}(M = 2) = \text{P}(3, 2, 2) + \text{P}(2, 3, 2) + \text{P}(2, 2, 3) + \text{P}(2, 2, 2)$
$= qpp + pqp + ppq + ppp$
$= 0.6 \times 0.4 \times 0.4 + 0.4 \times 0.6 \times 0.4 + 0.4 \times 0.4 \times 0.6 + 0.4 \times 0.4 \times 0.4$
$= 0.352$
$\text{P}(M = 10) = \text{P}(3, 3, 3) + \text{P}(3, 3, 2) + \text{P}(3, 2, 3) + \text{P}(2, 3, 3)$
$= qqq + qqp + qpq + pqq$
$= 0.6 \times 0.6 \times 0.6 + 0.6 \times 0.6 \times 0.4 + 0.6 \times 0.4 \times 0.6 + 0.4 \times 0.6 \times 0.6$
$= 0.352$

m	2	3
$\text{P}(M = m)$	0.352	0.648

e The median can only take the values 2 and 3.
Let $\text{P}(2) = p = 0.4$
Let $\text{P}(3) = q = 0.6$
$\text{P}(N = 2) = \text{P}(3, 2, 2) + \text{P}(2, 3, 2) + \text{P}(2, 2, 3) + \text{P}(2, 2, 2)$
$= qpp + pqp + ppq + ppp$
$= 0.6 \times 0.4 \times 0.4 + 0.4 \times 0.6 \times 0.4 + 0.4 \times 0.4 \times 0.6 + 0.4 \times 0.4 \times 0.4$
$= 0.352$
$\text{P}(N = 3) = \text{P}(3, 3, 3) + \text{P}(3, 3, 2) + \text{P}(3, 2, 3) + \text{P}(2, 3, 3)$
$= qqq + qqp + qpq + pqq$
$= 0.6 \times 0.6 \times 0.6 + 0.6 \times 0.6 \times 0.4 + 0.6 \times 0.4 \times 0.6 + 0.4 \times 0.6 \times 0.6$
$= 0.352$

n	2	3
$\text{P}(N = n)$	0.352	0.648

Chapter Review 6

1 a A list of all the patients on the surgery books.

b A patient.

2 a Any two from:
It would take too long.
It would cost too much.
It could be difficult to get hold of all members.

b A list of all members of the gym.

c A member of the gym.

3 a A sampling frame has to be some sort of list – it may not be possible to list a population.

b A sample is usually easier to do, quicker to do and not as costly as a census.
(Also, a census is not appropriate if the testing process would destroy each sampling unit.)

4 a A statistic is a quantity calculated solely from the observations of a sample.

b i is a statistic

ii is not a statistic as it depends on the value μ

5 a The light bulbs would all be destroyed.

b A light bulb.

6 a Any two from:
It is quicker to do.
It is cheaper to do.
It is easier to do.

b It can be biased OR it is subject to natural variations.

c A numbered list of all 400 call centre operators.

d A call-centre operator.

e Yes, because he is using only the values from a sample. There are no parameters.

7 Any two from:
It gives everyone's views.
It is unbiased.
It is easy to conduct a census when the population has only 10 members.

8 **a** and **b** are statistics, **c** and **d** are not statistics since they involve a population parameter.

9 a $\mathrm{E}(X)=\frac{1}{2}\times 5+\frac{1}{3}\times 10+\frac{1}{6}\times 20=\frac{55}{6}$

$\mathrm{Var}(X)=\frac{1}{2}\times 5^2+\frac{1}{3}\times 10^2+\frac{1}{6}\times 20^2-\left(\frac{55}{6}\right)^2=\frac{1025}{36}$

b (5, 5), (10, 10), (20, 20), (5, 10), (10, 5), (5, 20), (20, 5), (10, 20), (20, 10)

c The sample space diagram shows the total of two randomly chosen coins.

	Coin 1					
Coin 2	**5**	**5**	**5**	**10**	**10**	**20**
5	10	10	10	15	15	25
5	10	10	10	15	15	25
5	10	10	10	15	15	25
10	15	15	15	20	20	30
10	15	15	15	20	20	30
20	25	25	25	30	30	40

The sample space diagram shows the mean of two randomly chosen coins.

	Coin 1					
Coin 2	**5**	**5**	**5**	**10**	**10**	**20**
5	5	5	5	7.5	7.5	12.5
5	5	5	5	7.5	7.5	12.5
5	5	5	5	7.5	7.5	12.5
10	7.5	7.5	7.5	10	10	15
10	7.5	7.5	7.5	10	10	15
20	12.5	12.5	12.5	15	15	20

The sampling distribution for $\overline{Y}$ is shown in the table.

y	5	7.5	10	12.5	15	20
$\mathrm{P}(Y=y)$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{9}$	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{36}$

10 a (6, 6, 6), (6, 6, 10), (6, 10, 6), (10, 6, 6), (6, 10, 10), (10, 6, 10), (10, 10, 6), (10, 10, 10)

b The median can only take the values 6 and 10.
Let $\mathrm{P}(6) = p = 0.6$
Let $\mathrm{P}(10) = q = 0.4$

$\mathrm{P}(N = 6) = \mathrm{P}(6, 6, 6) + \mathrm{P}(6, 6, 10) + \mathrm{P}(6, 10, 6) + \mathrm{P}(10, 6, 6)$
$= ppp + ppq + pqp + qpp$
$= 0.6 \times 0.6 \times 0.6 + 0.6 \times 0.6 \times 0.4 + 0.6 \times 0.4 \times 0.6 + 0.4 \times 0.6 \times 0.6$
$= 0.648$

$\mathrm{P}(N = 10) = \mathrm{P}(6, 10, 10) + \mathrm{P}(10, 6, 10) + \mathrm{P}(10, 10, 6) + \mathrm{P}(10, 10, 10)$
$= pqq + qpq + qqp + qqq$
$= 0.6 \times 0.4 \times 0.4 + 0.4 \times 0.6 \times 0.4 + 0.4 \times 0.4 \times 0.6 + 0.4 \times 0.4 \times 0.4$
$= 0.352$

n	6	10
$\mathrm{P}(N = n)$	0.648	0.352

10 c The mode can only take the values 6 and 10.

Let P(6) = p = 0.6

Let P(10) = q = 0.4

$P(M = 6) = P(6, 6, 6) + P(6, 6, 10) + P(6, 10, 6) + P(10, 6, 6)$

$= ppp + ppq + pqp + qpp$

$= 0.6 \times 0.6 \times 0.6 + 0.6 \times 0.6 \times 0.4 + 0.6 \times 0.4 \times 0.6 + 0.4 \times 0.6 \times 0.6$

$= 0.648$

$P(M = 10) = P(6, 10, 10) + P(10, 6, 10) + P(10, 10, 6) + P(10, 10, 10)$

$= pqq + qpq + qqp + qqq$

$= 0.6 \times 0.4 \times 0.4 + 0.4 \times 0.6 \times 0.4 + 0.4 \times 0.4 \times 0.6 + 0.4 \times 0.4 \times 0.4$

$= 0.352$

m	3	2
$P(M = m)$	0.648	0.352

Challenge

a $$E(\bar{X}) = E\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right)$$

$$= \frac{1}{n}E(X_1 + X_2 + \ldots + X_n)$$

$$= \frac{1}{n}\left(E(X_1) + E(X_2) + \ldots + E(X_n)\right)$$

$$= \frac{1}{n}(\mu + \mu + \ldots + \mu)$$

$$= \frac{1}{n}(n\mu)$$

$$= \mu$$

b $$\mathrm{Var}(\bar{X}) = \mathrm{Var}\left(\frac{X_1 + X_2 + X_3}{3}\right)$$

$$= \frac{1}{9}\mathrm{Var}(X_1 + X_2 + X_3)$$

$$= \frac{1}{9}\left(\mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \mathrm{Var}(X_3)\right)$$

$$= \frac{1}{9}(\sigma^2 + \sigma^2 + \sigma^2)$$

$$= \frac{\sigma^2}{3}$$