



Edexcel AS Mathematics

8MA0/21 - Statistics

Friday 23 May 2025

Questions and Fully Worked Model Solutions

This booklet places the **question and the solution together**. The mathematics is set out using proper fractions, powers, roots and exact notation, with extra written explanation so that a student can see *why* each step is being taken.

Solutions have been checked against the Pearson Edexcel Summer 2025 mark scheme. The aim is a clear **full-mark model response**, not merely a list of final answers.

Contents

Question 1	2
Question 2	3
Question 3	5
Question 4	7
Question 5	8

www.mathswithmatt.co.uk

Question 1

Question 1

The random variable $X \sim B(20, 0.37)$.

- (a) Find $P(X = 8)$.
- (b) Find $P(X \leq 5)$.
- (c) Find $P(X \leq 5.7)$.

Model solution

(a) Using the binomial probability formula,

$$P(X = 8) = \binom{20}{8} (0.37)^8 (0.63)^{12} = 0.172968 \dots$$

$$P(X = 8) = 0.173 \text{ (3 s.f.)}$$

((a) 1)

(b)

$$P(X \leq 5) = 0.191004 \dots$$

$$P(X \leq 5) = 0.191 \text{ (3 s.f.)}$$

((b) 1)

(c) Since X is discrete and can take integer values only, $X \leq 5.7$ is equivalent to $X \leq 5$.

$$P(X \leq 5.7) = P(X \leq 5) = 0.191004 \dots$$

$$P(X \leq 5.7) = 0.191 \text{ (3 s.f.)}$$

((c) 1)

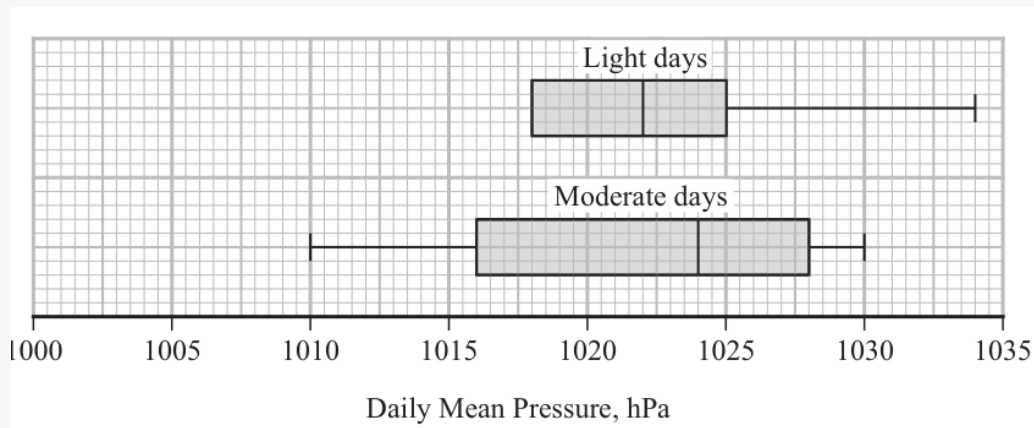
Question 2

Question 2

Jasper is investigating the relationship between Daily Mean Pressure and Daily Mean Windspeed (Beaufort conversion) for Perth in 2015, using the large data set. Treating the data for Perth in 2015 as the population, he uses all the data available.

(a) Write down the name given to this method of data collection.

All the Daily Mean Windspeed data for Perth in 2015 are classed as either Light or Moderate. Jasper draws a box plot for each set of data, but omits the lower tail for Light days.



The smallest three Daily Mean Pressures for Light days are 1007, 1009 and 1010 hPa. An outlier in the first quartile is any value more than $1.5 \times \text{IQR}$ below Q_1 .

(b) (i) Determine whether there are any outliers in the first quartile of Daily Mean Pressure for Light days.

(ii) Hence, complete the box plot of Daily Mean Pressure for Light days.

The box plot for Light days is based on data for 161 days.

(c) Using your knowledge of the large data set, estimate the number of Moderate days with a Daily Mean Pressure of 1028 hPa or higher in the data set for Perth in 2015.

Model solution

(a) Every member of the population is included.

Census.

((a) 1)

(b)(i) From the Light-days box plot, $Q_1 = 1018 \text{ hPa}$ and $Q_3 = 1025 \text{ hPa}$, so

$$\text{IQR} = 1025 - 1018 = 7 \text{ hPa}.$$

$$\text{Lower outlier boundary} = 1018 - 1.5(7) = 1007.5 \text{ hPa}.$$

Since $1007 < 1007.5$, 1007 hPa is an outlier. The values 1009 hPa and 1010 hPa are not outliers.

There is one lower outlier: 1007 hPa.

((b)(i) 2)

(b)(ii) The lower whisker must end at the smallest non-outlier, 1009 hPa, and the value 1007 hPa must be plotted separately as an outlier.

Lower whisker at 1009 hPa, with a single outlier plotted at 1007 hPa.

((b)(ii) 1)

(c) There are 184 Perth records in the large data set for 2015. Hence the number of Moderate days is

$$184 - 161 = 23.$$

The upper quartile of the Moderate-days box plot is 1028 hPa, so approximately one quarter of Moderate days have pressure at least 1028 hPa.

$$0.25(23) = 5.75.$$

Approximately 6 Moderate days.

((c) 2)

Question 3

Question 3

A bag contains red, green and white counters only. Elsa believes that their proportions are as follows.

Colour	red	green	white
Proportion	p	0.2	$4p$

Elsa selects 40 counters at random, one at a time with replacement.

Assuming Elsa's belief is true:

(a) Find the distribution of the number of red counters she selects.

Jayda believes that the true proportion of green counters is greater than 0.2. She takes a random sample of 40 counters, one at a time with replacement, and obtains 11 green counters.

(b) (i) Use a suitable test to assess Jayda's belief. State hypotheses clearly, use a 5% significance level, and state the p-value.

(ii) Find the acceptance region for the test in part (i).

Model solution

(a) The proportions sum to 1, so

$$p + 0.2 + 4p = 1, \quad 5p = 0.8, \quad p = 0.16.$$

Each selection is independent because the counter is replaced. Thus, if R is the number of red counters selected,

$$R \sim B(40, 0.16).$$

((a) 3)

(b)(i) Let π be the true proportion of green counters. Since Jayda's claim is that it is greater than 0.2, use a one-tailed upper test:

$$H_0 : \pi = 0.2, \quad H_1 : \pi > 0.2.$$

Under H_0 , if X is the number of green counters in the sample, then

$$X \sim B(40, 0.2).$$

The observed value is 11, giving

$$\text{p-value} = P(X \geq 11) = 1 - P(X \leq 10) = 1 - 0.8392 \dots = 0.1607 \dots$$

Since $0.1607 \dots > 0.05$, the result is not significant at the 5% level. Do not reject H_0 .

p-value = 0.161 (approximately 0.16); there is insufficient evidence to support Jayda's belief that the proportion of green counters is greater than 0.2.

((b)(i) 4)

(b)(ii) For an upper-tailed test, find the largest accepted value. Under H_0 ,

$$P(X \leq 11) = 0.9125 \dots < 0.95, \quad P(X \leq 12) = 0.9568 \dots > 0.95.$$

Therefore the critical region is $X \geq 13$, so the acceptance region is

$$0 \leq X \leq 12.$$

((b)(ii) 2)

www.mathswithmatt.co.uk

Question 4

Question 4

Giovanni believes there is a relationship between the mass of a car, m kg, and its fuel consumption, c miles per gallon (mpg). For a sample of 25 cars,

$$\sum (c - \bar{c})^2 = 394.$$

(a) Find the standard deviation of the fuel consumption of the 25 cars.

Using m as the explanatory variable, Giovanni creates the linear regression model

$$c = a + bm,$$

where a and b are constants. On average, the fuel consumption is 3.5 mpg lower for each additional 500 kg of mass, and the fuel consumption of a car of mass 1700 kg is 20 mpg.

(b) (i) Find b .

(ii) Find a .

Model solution

(a) The standard deviation for these 25 data values is

$$\sigma_c = \sqrt{\frac{\sum (c - \bar{c})^2}{25}} = \sqrt{\frac{394}{25}} = 3.9698 \dots$$

$$\sigma_c = 3.97 \text{ mpg (3 s.f.)}$$

((a) 1)

(b)(i) An increase of 500 kg decreases the predicted fuel consumption by 3.5 mpg. Hence

$$b = \frac{-3.5}{500} = -0.007.$$

$$b = -0.007 \text{ mpg kg}^{-1}.$$

((b)(i) 2)

(b)(ii) Substituting $m = 1700$ and $c = 20$ into the model gives

$$20 = a + (-0.007)(1700) = a - 11.9.$$

$$a = 31.9.$$

$$a = 31.9.$$

((b)(ii) 2)

Question 5

Question 5

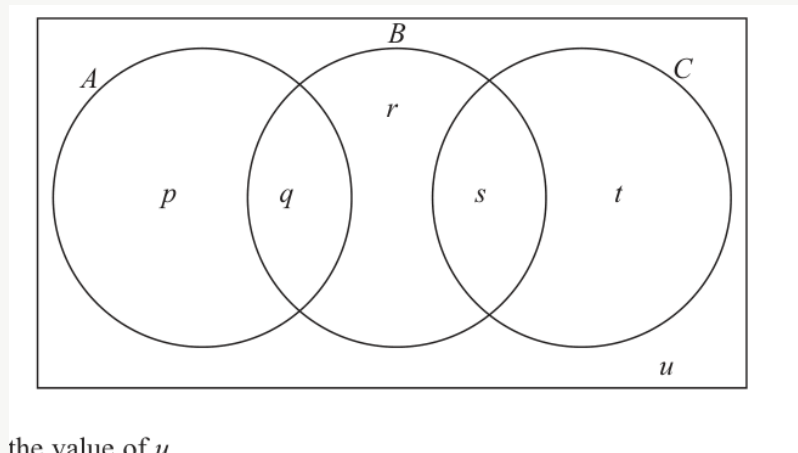
Events A, B and C satisfy

$$P(A \text{ or } C) = 0.55,$$

the probability that none occurs is 0.23, the probability that exactly two occur is 0.18, A and B are independent, and

$$P(B) = P(A) + 0.1.$$

The labelled Venn diagram is shown below.



the value of u

- Write down the value of u .
- Find the value of r .
- Find (i) p , (ii) q , (iii) s , and (iv) t .

Model solution

(a) The region outside all three events represents the probability that none occurs.

$$u = 0.23.$$

((a) 1)

(b) The event A or C consists of the regions p, q, s, t , so

$$p + q + s + t = 0.55.$$

All regions shown have total probability 1. Therefore

$$r + 0.55 + u = 1, \quad r = 1 - 0.55 - 0.23 = 0.22.$$

$$r = 0.22.$$

((b) 2)

(c) Exactly two events occur in the regions q and s , hence

$$q + s = 0.18.$$

Using $r = 0.22$,

$$P(B) = q + r + s = 0.18 + 0.22 = 0.40.$$

Since $P(B) = P(A) + 0.1$,

$$P(A) = 0.40 - 0.10 = 0.30.$$

Events A and B are independent, and their intersection is the region q . Thus

$$q = P(A \cap B) = P(A)P(B) = 0.30 \times 0.40 = 0.12.$$

$$s = 0.18 - 0.12 = 0.06.$$

Finally, $P(A) = p + q$, and $P(A \text{ or } C) = p + q + s + t$, so

$$p = 0.30 - 0.12 = 0.18, \quad t = 0.55 - (0.18 + 0.12 + 0.06) = 0.19.$$

(i) $p = 0.18$, (ii) $q = 0.12$, (iii) $s = 0.06$, (iv) $t = 0.19$.

((c) 4)