# A2 Statistics Practice Paper D
# 60 Marks

60 Minutes

1. The number of students doing A Level Maths in a school is summarised in the table below. The Head of Mathematics is going to take a stratified sample of size 20 to survey the students.

| | **Year 12** | **Year 13** |
|---|---|---|
| **Male** | 45 | 40 |
| **Female** | 35 | 40 |

a. State clearly what type of sample he should take and describe the sample. **(3)**

A factory produces components. Each component has a unique identity number and it is assumed that 2% of the components are faulty. On a particular day, a quality control manager wishes to take a random sample of 50 components.

b. Identify the sampling frame. **(1)**

The statistic $F$ represents the number of faulty components in the random sample of size 50.

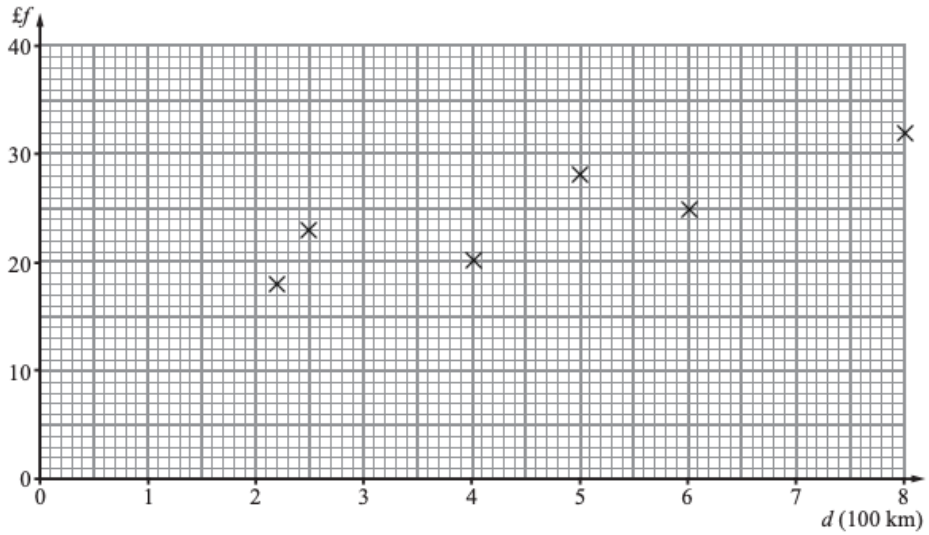
c. Specify the sampling distribution of $F$. **(2)**

**(Total marks: 6)**

2. A travel agent sells flights to different destinations from Beerow airport. The distance $d$, measured in 100 km, of the destination from the airport and the fare £$f$ are recorded for a random sample of 6 destinations.

| Destination | *A* | *B* | *C* | *D* | *E* | *F* |
|---|---|---|---|---|---|---|
| $d$ | 2.2 | 4.0 | 6.0 | 2.5 | 8.0 | 5.0 |
| $f$ | 18 | 20 | 25 | 23 | 32 | 28 |

[You may use $\Sigma d^2 = 152.09$ $\Sigma f^2 = 3686$ $\Sigma fd = 723.1$]



a. Explain why a linear regression model may be appropriate to describe the relationship between $f$ and $d$. **(1)**

b. Calculate $S_{dd}$ and $S_{fd}$ **(4)**

c. Calculate the equation of the regression line of $f$ on $d$ giving your answer in the form $f = a + bd$ **(4)**

d. Give an interpretation of the value of 2.03. **(1)**

Jane is planning her holiday and wishes to fly from Beerow airport to a destination $t$ km away. A rival travel agent charges 5p per km.

e. Find the range of values of $t$ for which the first travel agent is cheaper than the rival. **(2)**

**(Total marks: 12)**

3. The heights of adult females are normally distributed with mean 160 cm and standard deviation 8 cm.

a. Find the probability that a randomly selected adult female has a height greater than 170 cm. **(3)**

Any adult female whose height is greater than 170 cm is defined as tall.

b. An adult female is chosen at random. Given that she is tall, find the probability that she has a height greater than 180 cm.

**(4)**

Half of tall adult females have a height greater than $h$ cm.

c. Find the value of $h$. **(5)**

**(Total marks: 12)**

4. A bag contains a large number of 10p, 20p and 50p coins in the ratio 1: 2: 2

A random sample of 3 coins is taken from the bag.

Find the sampling distribution of the median of these samples.

**(Total marks: 7)**

5. The weight, in grams, of beans in a tin is normally distributed with mean $\mu$ and standard deviation 7.8

a. Given that 10% of tins contain less than 200 g, find the value of $\mu$. **(3)**

b. Find the percentage of tins that contain more than 225 g of beans. **(3)**

The machine settings are adjusted so that the weight, in grams, of beans in a tin is normally distributed with mean 205 and standard deviation $\sigma$.

c. Given that 98% of tins contain between 200 g and 210 g find the value of $\sigma$. **(4)**

**(Total marks: 10)**

6a. A shopkeeper knows, from past records, that 15% of customers buy an item from the display next to the till.

After a refurbishment of the shop, he takes a random sample of 30 customers and finds that only 1 customer has bought an item from the display next to the till.

Stating your hypotheses clearly, and using a 5% level of significance, test whether or not there has been a change in the proportion of customers buying an item from the display next to the till.

**(6)**

b. During the refurbishment a new sandwich display was installed. Before the refurbishment 20% of customers bought sandwiches. The shopkeeper claims that the proportion of customers buying sandwiches has now increased. He selects a random sample of 120 customers and finds that 31 of them have bought sandwiches.

Using a suitable approximation and stating your hypotheses clearly, test the shopkeeper's claim. Use a 10% level of significance.

**(7)**

**(Total marks: 14)**

---

**Total Marks for Paper: 60**

# Mark Scheme

| | | |
|---|---|---|
| 1a | Stratified sampling | **M1** |
| | <table><tr><th></th><th>Year 12</th><th>Year 13</th></tr><tr><td>**Male**</td><td>6</td><td>5</td></tr><tr><td>**Female**</td><td>4</td><td>5</td></tr></table> | **A1**<br>**A1** |
| 1b | List of ID numbers | **B1** |
| 1c | Binomial | **B1** |
| | $F \sim B(50, 0.02)$ | **B1** |

| | | |
|---|---|---|
| 2a | The points lie reasonably close to a straight line | **B1** |
| 2b | $\sum d = 27.7$<br>$\sum f = 146$ | |
| | $S_{dd} = 152.09 - \frac{27.7^2}{6}$<br>$= 24.208\ldots$<br>$= 24.2$ | **M1**<br>**A1** |
| | $S_{fd} = 723.1 - \frac{27.7 \times 146}{6} = 49.06$<br>$= 49.1$ | **A1** |
| 2c | $b = \frac{49.06\ldots}{24.208\ldots}$ | **M1** |
| | $b = 2.260\ldots$<br>$b = 2.03$ | **A1** |
| | $a = \frac{146}{6} - b \times \frac{27.7}{6}$ | **M1** |
| | $a = 14.97\ldots$<br>$f = 15.0 + 2.03d$ | **A1** |
| 2d | A flight costs £2.03 for every extra 100 km | |
| 2e | $15.0 + 2.03d < 5d$ | **M1** |
| | $t > 500 \sim 505$ | **A1** |

| | | |
|---|---|---|
| 3a | $P(H > 170) = P\left(Z > \frac{170-160}{8}\right)$ | **M1** |
| | $1 - 0.8994$ | **M1** |
| | $= 0.1056$ | **A1** |
| 3b | $P(H > 180) = P\left(z > \frac{180-160}{8}\right)$ | **M1** |
| | $= 0.0062$ | **A1** |
| | $[P(H > 180 \mid H > 170)] = \frac{0.0062}{0.1056}$ | **M1** |
| | $= 0.0587$ | **A1** |
| 3c | $P(H > h \mid H > 170) = 0.5$ | **M1** |
| | $P(H > h) = 0.5 \times 0.1056 = 0.0528$ | **A1** |
| | $\frac{h-160}{8} = 1.62$ | **M1**<br>**B1** |
| | $h = 173$ cm | **A1** |

4

**NB: If there is a fully correct table award full marks.**

| | | |
|---|---|---|
| P(10) = 0.2, P(20) = 0.4 and P(50) = 0.4 | B1: using P(10) = 0.2 ($p$) P(20) = 0.4($q$) and P(50) = 0.4($r$) may be seen in calculations or implied by a correct probability. | B1 |
| Median 10, 20, 50 | B1: three correct medians and no extras. | B1 |
| P(Median 10) = $0.2^3+3\times0.2^2\times0.4+3\times0.2^2\times0.4$ **or** $0.2^3+3\times0.2^2\times0.8$ | M1: allow if $(p+q+r)=1$ and use $p^3+3\times p^2\times q+3\times p^2\times r$ **or** $p^3+3\times p^2\times(q+r)$ look for $\frac{1}{125}+\frac{6}{125}+\frac{6}{125}$ | See below for how to award |
| P(Median 50) = $0.4^3+3\times0.4^2\times0.2+3\times0.4^2\times0.4$ **or** $0.4^3+3\times0.4^2\times0.6$ | M1: allow if $(p+q+r)=1$ and use $r^3+3\times r^2\times p+3\times r^2\times q$ **or** $r^3+3\times r^2\times(p+q)$ Look for $\frac{8}{125}+\frac{12}{125}+\frac{24}{125}$ | |
| P(Median 20) = $3\times0.2\times0.4^2+6\times0.2\times0.4\times0.4+0.4^3+3\times0.4^2\times0.4$ | M1: allow if $(p+q+r)=1$ and use $3\times p\times q^2+6\times p\times q\times r+q^3+3\times q^2\times r$ $\frac{12}{125}+\frac{24}{125}+\frac{8}{125}+\frac{24}{125}$ | |

**<u>How to award the M marks</u> – Allow the use of 1, 2 and 5 for the medians for the method marks**

**M1** any correct calculation (implied by correct answer) for P(m = 10) or P(m = 20) or P(m = 50)

**M1** any 2 correct calculations (implied by 2 correct answers) P(m = 10) or P(m = 20) or P(m = 50)

**M1** any 3 correct calculations (implied by 3 correct answers) for P(m = 10) and P(m = 20) and P(m = 50) or

3 probabilities that add up to 1 providing it is 1 – their 2 other calculated probabilities. Do **not** allow $\frac{1}{5}$ $\frac{2}{5}$ $\frac{2}{5}$

**NB** if they do not have a correct answer their working must be clear including the addition signs.

| median | 10 | 20 | 50 |
|---|---|---|---|
| | 0.104 | 0.544 | 0.352 |
| | Or $\frac{13}{125}$ | Or $\frac{68}{125}$ | Or $\frac{44}{125}$ |

A1: awrt any 1 correct
A2: awrt all 3 correct
These do not need to be in a table as long as the correct probablity is with the correct median(10, 20 & 50)
**NB: Do Not allow the use of 1,2 and 5 for the medians for the A marks**

A2

| | | |
|---|---|---|
| 5a | $\frac{200-\mu}{7.8} = -1.2816$ | **M1**<br>**B1** |
| | $\mu = 209.996 \ldots$<br>$\mu = 210$ | **A1** |
| 5b | $P(X > 225) = P\left(Z > \frac{225-210}{7.8}\right)$ | **M1** |
| | $P(Z > 1.92)$ | **A1** |
| | 1 – 0.9726<br>= 0.0274 | **A1** |
| 5c | $\frac{210-205}{\sigma} = 2.3263$ | **M1**<br>**B1** |
| | $\sigma = \frac{5}{2.3263}$ | **M1** |
| | $\sigma = 2.15$ | **A1** |

| | | |
|---|---|---|
| 6a | $H_0$: $p = 0.15$<br>$H_1$: $p \neq 0.15$ | **B1**<br>**B1** |
| | $X \sim B(30, 0.15)$ | **M1** |
| | $P(X \leq 1) = 0.0480$ | **A1** |
| | 0.0480 > 0.025<br>Not significant therefore do not reject $H_o$/not in critical region. | **M1** |
| | There is no evidence of a change in the proportion of customers buying an item from the display | **A1** |
| 6b | $H_0$: $p = 0.2$<br>$H_1$: $p > 0.2$ | **B1** |
| | Let S = the number who buy sandwiches, $S \sim B(120, 0.2)$<br>$S \approx W \sim N\left(24, \sqrt{19.2}^2\right)$ | **M1**<br>**A1** |
| | $P(S \geq 31) = P(W \geq 30.5)$ | **M1** |
| | $= P\left(Z > \frac{30.5-24}{\sqrt{19.2}}\right)$ | **M1** |
| | $= P(Z > 1.48 \ldots)$<br>= 1 – 0.9306 | **M1** |
| | = 0.0694 | **A1** |
| | Value is smaller than 0.10 so a significant result.<br><br>There is evidence that more customers are purchasing sandwiches or,<br>The shopkeepers claim is correct. | **B1** |