GCSE Stats Revision Paper 2

36 marks - 40 minutes (ET +10 minutes)

Higher Tier Formulae

You must not write on this page.

Anything you write on this page will gain NO credit.

$$\text{Skew} = \frac{3(\text{mean} - \text{median})}{\text{standard deviation}}$$

$$\text{Standard deviation} = \sqrt{\frac{1}{n}\sum(x - \bar{x})^2}$$

An alternative formula for standard deviation is

$$\text{standard deviation} = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

Spearman's rank correlation coefficient

$$r_s = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

Rates of change (e.g. Crude birth rate = $\frac{\text{number of births} \times 1000}{\text{total population}}$)

Solutions

1

Below is the table showing the population of Reigate and Bournemouth based on the 2021 cenus results.

Age Group	Bournemouth	Reigate
0-20 years	79,000	33,000
21-40 years	104,000	39,000
41-60 years	96,000	41,000
over 60 years	121,300	37,900

400300 150900

(a) Explain the meaning of a census. (1 mark)

All members of the population are surveyed.

(b) Using the information in the table calculate the standard population for Bournemouth and Reigate

Age Group	Bournemouth	Reigate
0-20 years	197.4	218.7
21-40 years	259.8	258.4
41-60 years	239.8	271.7
over 60 years	303	252.2

(4 marks)

$\frac{\text{stratum pop}^n}{\text{total pop}^n} \times 1000$

(c) The crude death rate for Bournemouth is 11.9 and the crude death rate for Reigate is 8.2. Using your standard population calculate the standardised death rate for Bournemouth and Reigate for the over 60s and compare these values. (3 marks)

B: $\frac{11.9 \times 303}{1000} = 3.61$

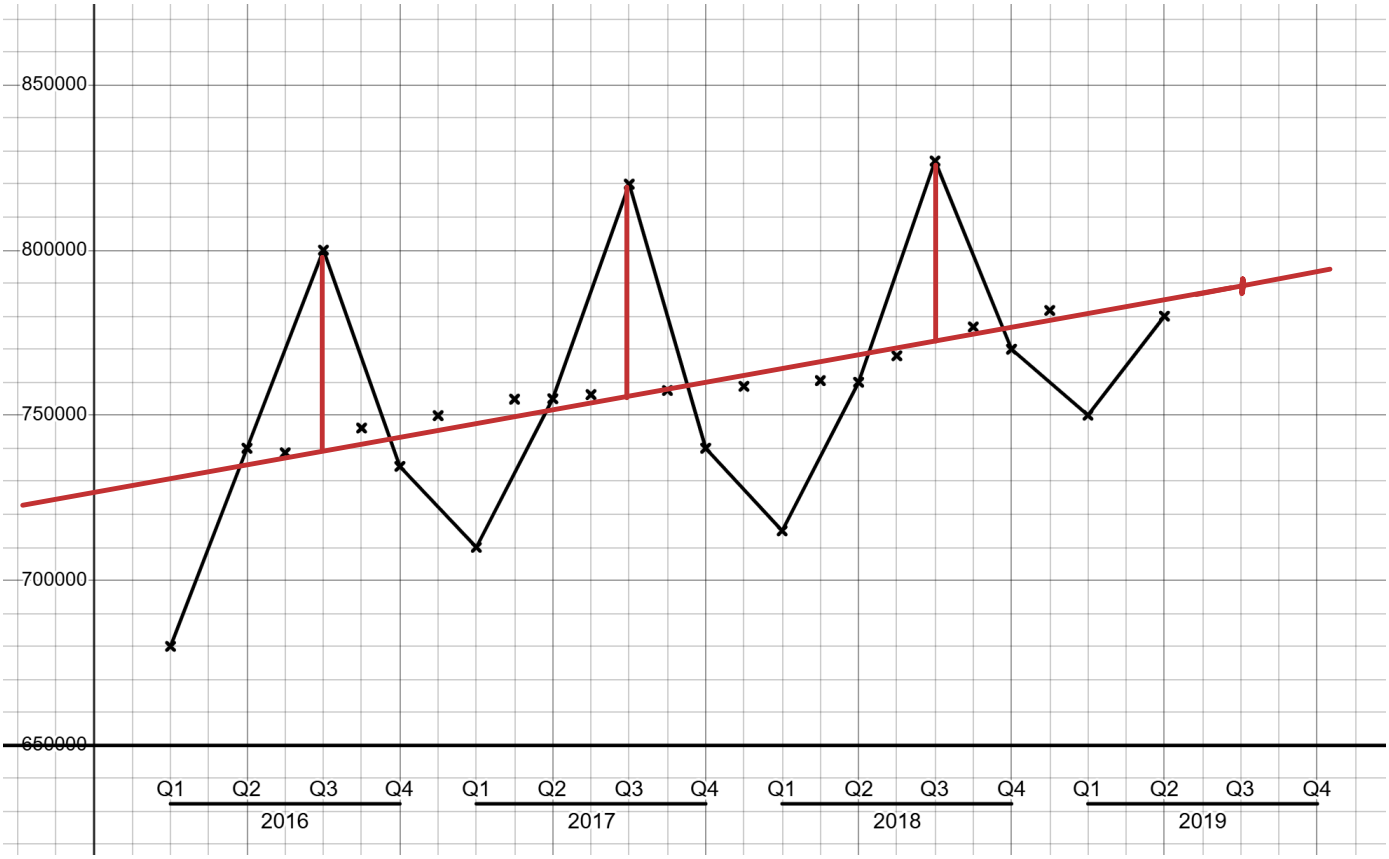
R: $\frac{8.2 \times 252.2}{1000} = 2.07$

The standardised death rate for over 60s for Bournemouth is 3.61 deaths per thousand, where as for Reigate over 60s it is 2.07 deaths per thousand.

STANDARISED RATE = $\frac{\text{CRUDE RATE}}{1000} \times$ STD POPN

2

This times series graph shows the number of visitors to the Tower of London each quarter from 2016 until 2019.



(a) Describe the trend shown by the time series graph. Increasing (1 mark)

(b) The moving averages have been calculated and plotted for you.
Draw in the trend line. (1 mark)

(c) Calculate the seasonal variation for quarter 3. (2 marks)

$$\frac{60000 + 65000 + 50000}{3} = \frac{175000}{3}$$

$$= 58300$$

(d) Calculate an estimate for the numbers of visitors (to the nearest 1000) in quarter 3 of 2019.

(3 marks)

$$790000 + 58300$$

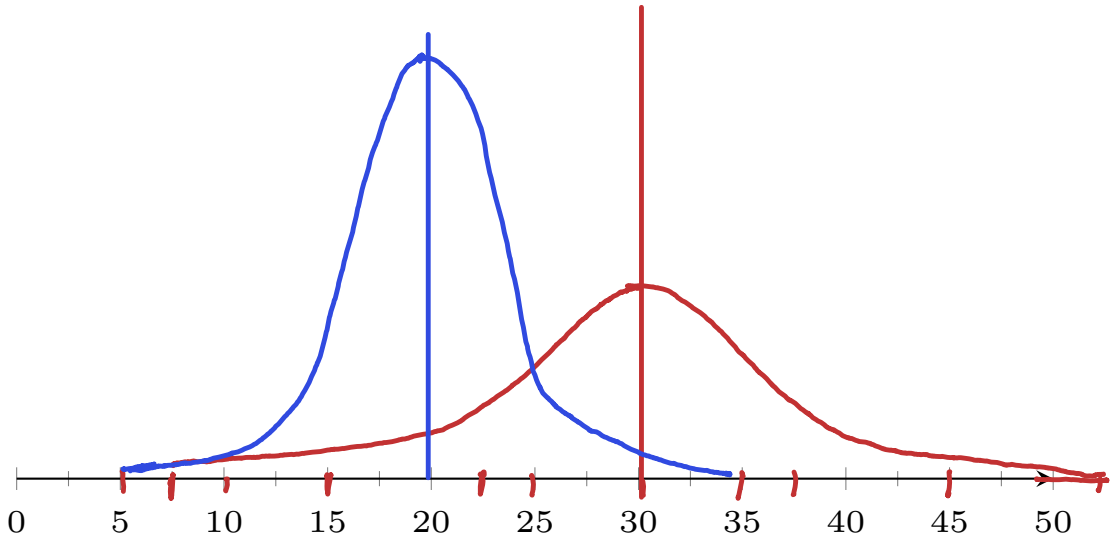
$$= 848000$$

3

On the same set of axes sketch the normal distributions A and B shown in the table

	A	B
Mean	30	20
Standard Deviation	7·5	5

(4 marks)



4

A heatlth study into drinking and heart disease had 189000 female participants.
Here is the data of the particupants drinking habits and whether they developed heart disease with in a year of the study.

Group	No Heart Disease	Heart Disease
Heavy Drinker	36372	1,398
Moderate Drinker	74673	901
Light/No Drinker	74847	809
	185892	3108

37770
755747 } 151230
75656

(a) Calculate the absolute risk of a female developing heart disease according to this study.

(3 mark)

$$\frac{3108}{189000} = 0.0164$$

(b) Calculate the relative risk of a female developing heart disease if they are a heavy drinker compare to if they are not a heavy drinker. (3 marks)

$$\frac{\text{Risk those in group}}{\text{Risk those not in group}} = \frac{\frac{1398}{37770}}{\frac{1710}{151230}}$$

$$= 3.27$$

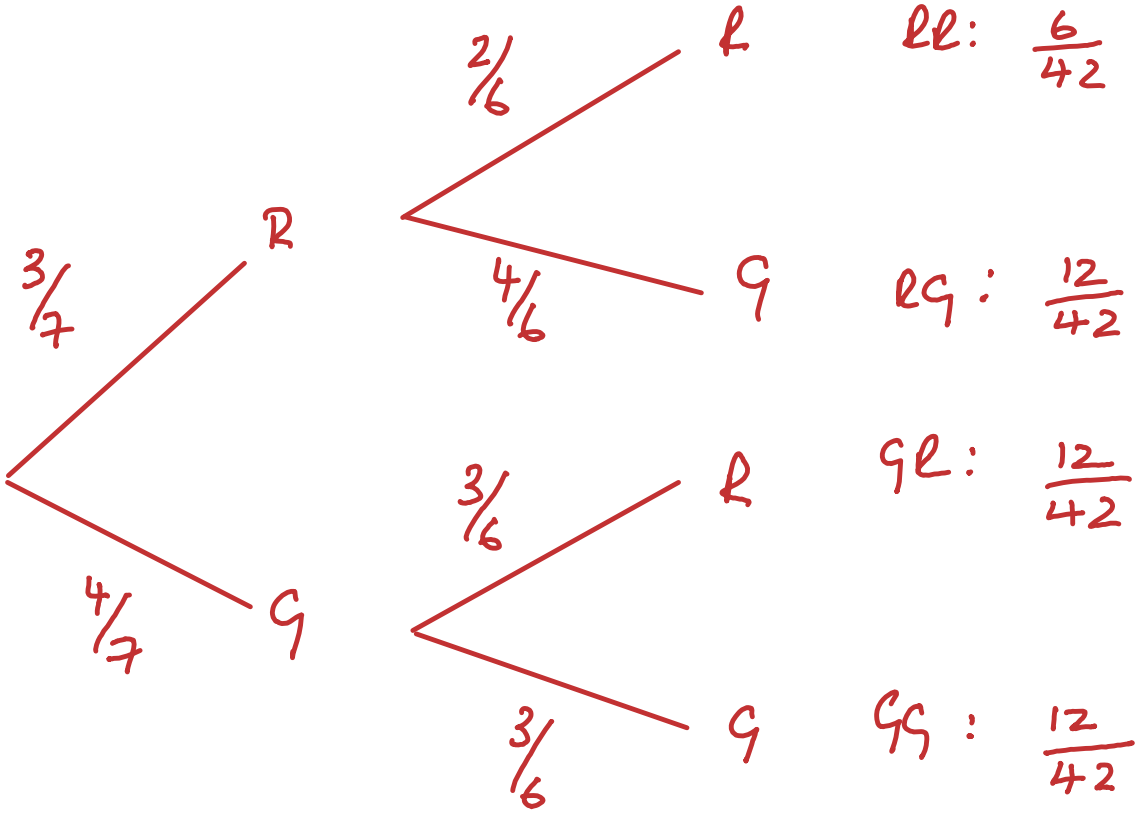
(c) Explain the meaning of your answer to (b) in context. (1 mark)

A heavy drinker is over 3 times more likely to develop heart disease than a non heavy drinker.

5

A bag contains 4 green beads and 3 red beads.
Two beads are drawn at random without replacement.
X is the number of green beads selected.

Work out the probability distribution for the X. (3 marks)



x	0	1	2
$P(X=x)$	$\frac{6}{42}$	$\frac{24}{42}$	$\frac{12}{42}$

6

John wants to investigate what proportion of students at his school have cheated on test.

(a) Explain why John could use the random response technique to provide him with valid data.

(3 mark)

- He is asking a sensitive question.
- Random response allows people to tell the truth without revealing they have told the truth.
- People are more likely to tell the truth.

(b) John randomly chooses one tutor group of size 30 in each of Years 7 - 10.
What is the name of the sampling method John has used? (1 mark)

Cluster.

John gives the following question to each student in his sample.

Flip a coin. If you get a head, tick box A. If you get a tail, answer the question: Have you ever cheated on a test? A. Yes B. No

(c) The data he collects from the 120 students is as follows:
A: 89
B: 31

Calculate an estimate for the percentage of students who have cheated in a test in John's school. (3marks)

$\frac{120}{2} = 60$

$89 - 60 = 29$

$\frac{29}{60} = 0.483$

48.3%