

Topic Test Summer 2022

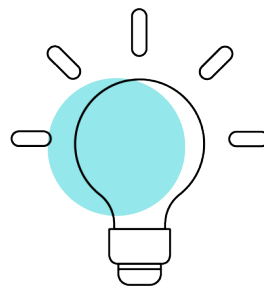
Pearson Edexcel GCE Mathematics (9MA0)

Paper 3 – Statistics

Topic 2: Measures of central tendency and variation

Contents

<u>General guidance to Topic Tests</u>	<u>3</u>
<u>Revise Revision Guide content coverage</u>	<u>4</u>
<u>Questions</u>	<u>5</u>
<u>Mark Scheme</u>	<u>19</u>



Mark Scheme

Question T2_Q1

Qu 4	Scheme	Marks	AO
(a)	Convenience <u>or</u> opportunity [sampling]	B1 (1)	1.2
(b)	Quota [sampling] e.g. Take 4 people every 10 minutes	B1 B1 (2)	1.1a 1.1b
(c)	Census	B1 (1)	1.2
(d)	[58 – 26 =] 32 (min)	B1 (1)	1.1b
(e)	$\mu = \frac{4133}{95} = 43.505263\dots$ $\sigma_x = \sqrt{\frac{202\,294}{95} - \mu^2} = \sqrt{236.7026\dots}$ $= 15.385\dots$ awrt 15.4 (min)	B1 M1 A1 (3)	1.1b 1.1b 1.1b
(f)	There are outliers in the data (or data is skew) which will affect mean and sd Therefore use median and IQR	B1 dB1 (2)	2.4 2.4
(g)	Value of 20, LQ at 26 and outliers will not change <u>or</u> state that median and upper quartile are the values that <u>do</u> change <u>More values now below 40 than above</u> so Q_2 <u>or</u> Q_3 will change and be lower Both Q_2 <u>and</u> Q_3 will be lower	B1 M1 A1 (3)	1.1b 2.1 2.4
(13 marks)			
Notes			
(b)	1 st B1 for quota (sampling) mentioned (“Stratified” or “systematic” or “random” are B0B0) 2 nd B1 for a description of how such a system might work, requires suitable strata or categories e.g. time slots, departments, gender, age groups, distance travelled etc Suggestion of randomness is B0		
(e)	B1 for a correct mean (awrt 43.5) M1 for a correct expression for the sd (including $\sqrt{}$)ft their mean A1 for awrt 15.4 (Allow $s = 15.4667\dots$ awrt 15.5)		
(f)	1 st B1 for acknowledging <u>outliers</u> or <u>skewness</u> are a problem for <u>mean and sd</u> “extreme values”/“anomalies” OK May be implied by saying median and IQR not affected by.. We need to see mention of “outliers”, “skewness” and the problem so “data is skewed so use median and IQR” is B0 unless mention that they are not affected by extreme values <u>or</u> mean and standard deviation can be “inflated” by the positive skew etc 2 nd dB1 dep on 1 st B1 for therefore choosing <u>median and IQR</u>		
(g)	B1 for identifying 2 of these 3 groups of unchanged values or stating only Q_2 and Q_3 change M1 for <u>explaining</u> that median or UQ should be lower. E.g. the 2 values have moved to below 40 (or 58) and therefore more than 50% below 40 or (more than 75% below 58) <u>or</u> an argument to show that the other 3 values are the same. (o.e.) Allow arrows on box plot provided statement in words about increased % below 40 or 58 etc A1 for stating median <u>and</u> UQ are both lower with clear evidence of M1 scored [If lots of values on 40 then median might not change but, since two values <u>do</u> change then UQ would change. If this meant that 92 became an outlier then we would have a new value for upper whisker and an extra outlier so effectively 3 values are altered. So median changes]		

Question T2_Q2

Question	Scheme		Marks	AOs
2(a)	IQR = 26.6 – 19.4 [= 7.2]		B1	2.1
	19.4 – 1.5 × ‘7.2’ [= 8.6] or 26.6 + 1.5 × ‘7.2’ [= 37.4]		M1	1.1b
	Plotting one upper whisker to 32.5 and one lower whisker to 8.6 or 9.1		A1	1.1b
	Plotting 7.6 and 8.1 as the only two outliers		A1	1.1b
			(4)	
(b)	October (since it is the month with the coldest temperatures between May and October in Beijing)		B1	2.4
			(1)	
(c)	$[\sigma = \sqrt{\frac{4952.906}{184}}$ or e.g. $[\sigma = \sqrt{\frac{S_{xx}}{n}} = 5.188...$ [=5.19*]		B1cso*	1.1b
			(1)	
(d)	$z = (\pm) 1.28(16)$	$[P_{90} =]29.251...$ or $[P_{10} =]15.948...$	B1	3.1b
	$2 \times 1.2816 \times 5.19$	‘29.251...’ – ‘15.948...’	M1	1.1b
	= awrt 13.3		A1	1.1b
			(3)	
(e)	Daily mean <u>wind speed</u> /Beaufort conversion since it is <u>qualitative</u>		B1	2.4
	<u>Rainfall</u> since it is not symmetric/lots of days with 0 rainfall		B1	2.4
			(2)	
(11 marks)				
Notes				
(a)	B1: for a correct calculation for the IQR (implied by 10.8 or 8.6 or 37.4 seen)			
	M1: for a complete method for either lower outlier limit or upper outlier limit (allow ft on their IQR) (may be implied by the 1 st A1 or a lower whisker at 8.6)			
	A1: both whiskers plotted correctly (allow ½ square tolerance)			
	A1: only two outliers plotted, 7.6 and 8.1 (must be disconnected from whisker)			
	NOTE: A fully correct box plot with no incorrect working scores 4/4			
(c)	B1cso*: Correct expression with square root or correct formula and 5.188 or better			
	Allow a complete correct method finding $\sum x^2 =$ awrt 98720 and $\sigma = \sqrt{\frac{98715.9...}{184} - \left(\frac{4153.6}{184}\right)^2}$			
(d)	B1: Identifying z-value for 10th or 90th percentile (allow awrt (±) 1.28) or for identifying $[P_{90} =]29.251...$ (awrt 29.3) or $[P_{10} =]15.948...$ (awrt 15.9) (This may be implied by a correct answer awrt 13.3)			
	M1: for $2 \times z \times 5.19$ where $1 < z < 2$ or for their $P_{90} - P_{10}$ where $25 < P_{90} < 35$ and $10 < P_{10} < 20$			
	A1: awrt 13.3			
(e)	B1: for one variable identified and a correct supporting reason			
	B1: for two variables identified and a correct supporting reason for each			
	Allow any two of the following:			
	• <u>Wind speed</u> /Beaufort since the data is <u>non-numeric</u> (o.e.). They need not mention Beaufort provided there is a description of the data as non-numeric (Do not allow wind direction/wind gust)			
	• <u>Rainfall</u> as not symmetric/is skewed/is not bell shaped/lots of 0s /many days with no rain/mean≠mode or median			
	• <u>Date</u> since each data value appears once/it is uniformly distributed			
	• Daily mean <u>pressure</u> since it is not symmetric/is skewed/not bell shaped			
	• Daily mean <u>wind speed</u> since it is not symmetric/is skewed/not bell shaped			
Do not allow ‘not continuous’ or ‘discrete’ as a supporting reason.				
Ignore extraneous non-contradicting statements				

Question T2_Q3

Qu 3	Scheme	Marks	AO
(a)	$[68 - 7 =]$ <u>61</u> (only)	B1 (1)	1.1b
(b)	$[25 - 14] =$ <u>11</u>	B1 (1)	1.1b
(c)	$\left[\mu \text{ or } \bar{x} = \frac{607.5}{27} = \right] =$ <u>22.5</u>	B1 (1)	1.1b
(d)	$\sigma = \sqrt{\frac{17\,623.25}{27} - "22.5"{}^2}$ <u>or</u> $\sqrt{146.4629...}$ = 12.10218... awrt <u>12.1</u>	M1 A1 (2)	1.1b 1.1b
(e)	$\mu + 3\sigma = "22.5" + 3 \times "12.1..." =$ awrt 59 so only <u>one</u> outlier	B1ft (1)	1.1b
(f)	Median increases implies that both values must be > 20 Mean is the same means that $a + b = 45$ So possible values are: e.g. $b = 21$ and $a = 24$ (o.e.)	M1 M1 A1 (3)	3.1b 1.1b 2.2b
(g)	Both values will be less than 1 standard deviation from the mean and so the standard deviation of all 29 values will be smaller	B1 (1)	2.4
		(10 marks)	
Notes			
(a)	B1 for correctly interpreting the box plot to find the range (more than 1 answer is B0)		
(b)	B1 for correct understanding of IQR and answer of 11		
(c)	B1 for 22.5 only (or exact equivalent such as $\frac{45}{2}$). Allow 22 mins and 30 secs.		
(d)	M1 for a correct expression including square root. Allow $\sqrt{146}$ or better. Ft their mean A1 for awrt 12.1 NB Allow use of $s = 12.3327...$ or awrt 12.3		
(e)	B1ft for a correct calculation or value based on their μ and σ and compatible conclusion		
(f)	1 st M1 Correct start to the problem and a correct statement about the values based on median Allow if their final two values are both > 20 2 nd M1 for a correct explanation leading to equation $a + b = 45$ (o.e. e.g. equidistant from mean) Allow if their final two values sum to 45 A1 for a correct pair of values (both > 20 with a sum of 45) and at least some attempt to explain how their values satisfy at least one of the conditions (both > 20 <u>or</u> $a + b = 45$). Ignore $a =$ or $b =$ labels NB The values for a and b do not need to be integers.		
(g)	B1 for a correct explanation. Must mention that both values are less than 1 sd (ft their answer to (d)) from the mean		

Question T2_Q4

Qu 3	Scheme	Marks	AO
(a)	Hectopascal <u>or</u> hPa	B1 (1)	1.2
(b)	$\bar{x} = \bar{y} + 1010$ <u>or</u> $\frac{214}{30} + 1010$ $= 1017.1333\dots$ awrt 1017	M1 A1 (2)	1.1b 1.1b
(c)	$\sigma_x = \sigma_y$ (or statement that standard deviation is not affected by this type of coding) $[\sigma_y =] \sqrt{\frac{5912}{30} - ("7.13[33\dots]")^2}$ <u>or</u> $\sqrt{146.1822\dots}$ $= 12.0905\dots$ awrt 12.1	M1 M1 A1 (3)	3.1b 1.1b 1.1b
(d)	High pressure (since approx. mean + sd) so clockwise Locations are (from North to South): Leuchars, Heathrow, Hurn Wind direction is direction wind blows <u>from</u> So: Heathrow (NE) Hurn (E) Leuchars (W)	B1 B1 (2)	2.4 2.2a
(8 marks)			
Notes			
FYI	1 hPa = 100 Pa; 10hPa = 1 kPa; 1Pa = 1 Nm ⁻²		
(a)	B1 for “hectopascal” <u>or</u> hPa (condone pascals, allow millibars <u>or</u> mb) o.e. Do NOT allow kPa <u>or</u> kilopascals <u>or</u> Pa on its own		
(b)	M1 for a strategy to find $\bar{x}$ Allow an attempt to find $\sum x$ that gets as far as $\sum x = \sum y + 30 \times 1010 [= 30\,514]$ A1 for awrt 1017 (accept 1020) [Ignore incorrect units]		
(c)	1 st M1 for an overall strategy using the fact $\sigma_x = \sigma_y$ (can be implied by correct <u>final</u> ans) <u>or</u> for $\sum x = 30\,514$ and $\sum x^2 = 31\,041\,192$ (both seen and correct) 2 nd M1 for a correct expression (with $\sqrt{}$) (ft their $\bar{y}$ to 3sf) allow awrt 146 for 146.1822.. <u>or</u> for correct expression in x can ft their $\sum x > 30\,000$ or their answer to (b) A1 (dep on 2 nd M1) for awrt 12.1 [Ignore incorrect units] Final ans of awrt 12.1 scores 3/3 but if they then adjust for x e.g. add 1010 (M0M1A1)		
Final answer	(d) 1 st B1 for at least one of these reasons (these 2 lines) clearly stated (may see diagram) Need “high pressure” and “clockwise” to score on 1 st line Contradictory statements B0 e.g. correct N~S list but say “anticlockwise” 2 nd B1 (indep of 1 st B1) for deducing the 3 correct directions either in the table or stated as above If the answers in table and text are different we take the table (as question says)		