A Level Statistics

AQA Past Exam Questions

Solutions

TOPIC: Sam

Some questions form the old AQA syllabus have been omitted as they require the use of the Random Number Tables which are no longer given in the formula booklet or used.

Some questions are also outdated with the advances in the use of technology

Candidates may use any calculator allowed by Pearson regulations. Calculators must not have retrievable mathematical formulae stored in them.

Instructions

• Use **black** ink or ball-point pen.

• If pencil is used for diagrams/sketches/graphs it must be dark (HB or B). Coloured pencils and highlighter pens must not be used.

• **Fill in the boxes** at the top of this page with your name, centre number and candidate number.

• Answer **all** questions and ensure that your answers to parts of questions are clearly labelled.

• Answer the questions ***on paper***

• You should show sufficient working to make your methods clear. Answers without working may not gain full credit.

• Unless otherwise stated, statistical tests should be carried out at the 5% significance level.

• When a calculator is used, the answer should be given to three significant figures unless otherwise stated.

Information

• ***You may use the*** booklet 'Statistical Formulae and Tables'

• There are **8** questions in this question paper. The total mark for this paper is ***127***

• The marks for **each** question are shown in brackets – use this as a guide as to how much time to spend on each question.

Advice

• Read each question carefully before you start to answer it.

• Try to answer every question.

• Check your answers if you have time at the end.

• If you change your mind about an answer, cross it out and put your new answer and any working underneath.

• Check your answers if you have time at the end.

AQA_JUNE_2015_6

Q6	Solution	Marks	Total	Comments
(a) (i)	Cluster sampling	B1		
(ii)	Because the list is alphabetical by family name so the sample may contain several members of the same family	E1		Any indication of the problems arising because the list is alphabetic
			2	
(b)	Use random numbers to select a patient between 1 and 80 Select every 80th patient after that.	M1 A1		Allow even if method of random selection is not given SC If M0 then "every 80th" gains B1
			2	
(c)(i)	Stratified sampling	B1		
(ii)	737 ÷ 3200 × 40 (= 9.2125) = 9	M1 A1		Must be integer
			3	
(d)(i)	2617 + 1 = 2618	M1 A1		
(ii)	Otherwise those numbered 1 to 400 would have a greater chance of being chosen than other numbers	E1		OE, being generous on details throughout part (d)
(iii)	Otherwise remainder 0 would not have a corresponding patient, Or otherwise patient 3200 could not be chosen	E1		Or random number 0000, 3200, 6400 stated For either of these
(iv)	Rejecting/ignoring any repeats	E1		
			5	

AQA_JUNE_2012_4a

Q	Solution	Marks	Total	Comments
4(a)(i)	Cluster sampling	B1	1	
(ii)	Hypothesis test needs random sample	E1	1	Reference to sample

AQA_JAN_2012_6

Q	Solution	Marks	Total	Comments
6(a)(i)	Use 3-figure random numbers	E1		
	Reject repeats, 000 and numbers > 500	E1		Condone not mentioning 000
	Continue until 50 numbers generated.	E1		
	Use the numbers to identify the animals from the stock book.	E1	4	If candidate uses 0 to 499 they must relate to stock number for this mark
(ii)	The random sample may not include any goats (or too many).	E1	1	For showing appreciation that number of goats may be disproportionate
(b)(i)	Systematic.	B1	1	
(ii)	Not random	B1		
	Not every group of 50 can be chosen (Eg if 7 then not 8).	E1	2	
(c)(i)	He decides how many of each type to test (Eg.33 sheep, 16 cattle and 1 goat)	M1		Not necessarily proportionately stratified.
	Then he tests **any** 33 sheep, 16 cattle and 1 goat that he finds.	E1	2	Consistent with above.
(ii)	Convenience **Or** Guarantees at least one of each type of animal. **Or** Gives correct proportions	E1	1	If stratified in (i)
(iii)	The sample may be **biased** – he might only test the slower animals.	E1	1	OE Not simply 'Not random' – must say why this is a disadvantage
	Total		**12**	

AQA_JAN_2013_6

6(a)	Stratified	B1	**1**	
(b)	Eg. No complete list of customers. No contact details for customers. Very time consuming Very expensive May not give desired proportions People would not want to be delayed in a **fast** food outlet	E3,2,1	**3**	Any three comments addressing different aspects. Expense and time count as separate points.
(c)(i)	Convenient. They just go locally and question customers until they have reached the numbers required.	E1		Convenience.
(ii)	Not representative. Other parts of the country may have different views.	E1	**2**	Not representative.
(d)	Eg. A small number of outlets. Randomly selected Or selected for spread of size etc Decide on a quota for each type of customer at the chosen outlets. Select the customers who will fulfil the quotas trying to avoid bias	E1 E1 E1 (E1) (E1)	**3**	Small number Random (outlets) **or** balanced Quota Avoid bias (customers) (Accept 'randomly select customers') Any additional valid point **Maximum of 3 marks**
		Total	**9**	

AQA_JUNE_2014_5

Q	Solution	Marks	Total	Comments
5(a)	Total for school = 750	B1		
	Bronwyn needs $^{50}/_{750} = {}^{1}/_{15}$ of population	M1		Possibly implied
	Attempt to divide each cell by 15	m1		
	Integer answers			
	Boys 4 5 5 6 6	A1		At least two ≠ 5 seen
	Girls 4 5 5 5 5			
			4	
(b)(i)	All the boys followed by all the girls	B1		Or vice versa
(ii)	Choose a number between 1 and 15 at random	B1		
	using random numbers, calculator, etc	B1		Any valid method suggested, dep on previous B1
	Select every fifteenth pupil after that	B1		
			4	
(c)	Advantage – does not need to find particular pupils, quicker or easier.	E1		Or similar, must be in context
	Disadvantage – groups arriving together are likely to have travelled together.	E1		Or similar, must be in context
			2	

AQA_JUNE_2017_7ab

Q	Solution	Marks	Total	Comments
(a)(i)	Stratified sampling	B1	1	
(ii)	46 men and 34 women	M1		PI by other figures
	23 men under 30 years and 23 older men			At least 2 out of 4 values correct and labelled A1, other 2 A1
	19 women under 30 years and 15 older women	A2, 1	3	
(b)(i)	Opportunity or cluster sampling	B1	1	
(ii)	Advantage: eg. Quicker and/or easier for James	E1		Or similar
	Disadvantage: eg. Will be biased towards members who use the cafeteria			If bias is mentioned, must say in what way, to get the mark.
	Or Opinions may be influenced by other members of the group.	E1		Or similar
	Or Difficulty of choosing groups to add up exactly to 80		2	

7 (a)	The list does not tell him the sex or age of the participants ….	E1		The problem with the list
	so he cannot pick a sample in the right proportions	E1	**2**	The consequence for the sample
	Alternative for one mark The list does not give any information about whether the runner exercises regularly	(E1)		
(b)(i)	Using random number tables or function select 4-digit random number	B1		Must specify 4-digit
	ignoring (0000 and) >2000 and repeats	B1		Condone not mentioning 0000
	continue until 40 have been selected (and measure those runners)	B1		
(b)(ii)	eg. difficult to find the numbered people at the end of the run some of the numbers chosen for his sample may belong to people who do not finish the run.	E1	**4**	Must be in this context and practical, not statistical
(c)(i)	Randomly choose one of the first 50 to finish	B1		Must specify random
	Choose every 50th person (after that) to be in the sample	B1		Indep of previous B1
(c)(ii)	Advantage: Eg. Should space out the people who need to be measured. Should give a spread of fitness levels.	E1		Must be in this context
	Disadvantage: Eg. May get a large number crossing the line close together – hard to manage	E1	**4**	Must be in this context
(d)(i)	Cluster sampling	B1		
(d)(ii)	Advantage: Eg. Easy to spot who is in the sample. Can give them instructions before the run	E1		Must be in this context
	Disadvantage: Eg. May be very unrepresentative, as one charity may be all the same age, all the same sex, all the same1 fitness level.	E1	**3**	Must be in this context
		Total	**13**	

AQA_JAN_2007_7

7(a)	Number drivers 000 to 619	E1		number drivers 000 to 619 or mechanics/clerical staff/managers
	Select 3 digit random numbers	E1		select 3 digit random numbers
	Ignore repeats and >619 Continue until 62 numbers obtained	E1 E1		ignore repeats ignore >619 (must be consistent with numbering)
	Select corresponding drivers	E1		idea of stratified sample
	Similarly select 12 mechanics,13 clerical staff, 80 managers	E1	6	12,13,8 or explanation why not necessarily so allow max 3 for random sample
(b)	Number all employees 000 to 949	E1		number 000 to 949
	Choose a random digit between 0 and 9	E1		choose a random digit between 0 and 9
	Select every 10th employee e.g. if 7 picked select 007,017.....947	E1	3	select every 10th employee
(c)	No point in stratifying by employment categories if no difference between categories	E1		
	Would be worth stratifying by sex	E1		
	Systematic sample would not ensure a fair representation of sexes (unless men numbered together and women numbered together)	E1	3	a mark for any sensible point - max 3
	Total		**12**	

AQA_JAN_2008_8

8(a)	Number staff 000 to 819	E1		Valid numbering
	Select 3-digit random numbers Ignore >819	E1		3-digit random numbers
	Ignore repeats	E1		ignore >819 and repeats
	Continue until 25 selected and choose corresponding staff	E1	4	continue until 25 selected
(b)(i)	Permit holder/waiting list/other	B1		permit holder status
	male/female full-time/part-time etc	B1	2	any other sensible strata
(ii)	(A) Choose a digit between 1 and 8 at random. Pick this space and every 8th thereafter. e.g. 3,11,19.....187,195	E1 E1		idea of systematic sampling correct method including "every 8th"
	(B) Easy and quick	B1		easy - or other valid advantage
	(C) Excludes anyone without a permit, favours those who usually arrive early etc	E1 E1	5	any reasonable source of possible bias any different reasonable source of possible bias
	Total		**11**	

AQA_JAN_2010_6

6(a)	Number examiners 000 to 399	B1 E1		400 examiners - may be implied valid numbering
	Select 3 digit random numbers	E1		select 3 digit random numbers
	Ignore repeats and greater than 399 Continue until 40 selected choose corresponding examiners	E1 E1	5	ignore repeats and out of range select 40 and choose corresponding examiners
(b)(i)	Cluster sampling	B1	1	cluster sampling
(ii)	More geographically localised - less travelling	E1	1	less travelling
(iii)	Views will differ between regions e.g. examiners from South East likely to prefer London and examiners from North West likely to prefer Manchester	E1 E1	2	views likely to be more homogeneous in context
(iv)	No - examiners from regions with small number of examiners e.g. North West more likely to be selected than those from regions with a large number of examiners.	B1 E1	2	no explanation allow B1 for no, examiners in regions not chosen have no chance
(c)	(i) (b)(ii) no longer valid since no travelling required (ii) (b)(iii) still valid - using telephone email does not affect it.	E1 E1	2	no travelling needed views will still differ between regions allow E1 for (i) no (ii) yes without explanation
	Total		**13**	

AQA_JAN_2011_5

5(a)	Number medical staff from 000 to 389. Choose 3-digit random numbers. Ignore repeats and > 389 Continue until 39 obtained. Choose corresponding medical staff. In the same way select 22 ancillary staff, 14 administrative staff and 7 managers.	E1 E1 E1 E1 E1 E1	6	E1 valid numbering of one strata E1 3-digit random numbers E1 ignore repeats E1 ignore > 389 *or equivalent for another strata* E1 similarly for other strata E1 39,22,14,7
(b)(i)	Number medical staff 000 to 389, ancillary staff 390 to 609, administrative staff 610 to 749 and managers 750 to 819 Select a random number between 000 and 027. Choose this number and every 8th number thereafter until 100 have been selected. Choose corresponding staff.	E1 E1 E1 E1		E1 valid numbering E1 choose random starting point (not necessarily in range 000 to 027) E1 idea of systematic sampling E1 choose every 8th
(ii)	Because 820 is not exactly divisible by 100. (In the sample described above numbers 000 to 027 have different chances of being selected the rest have a 1 in 8 chance.)	E1	5	E1 820 not exactly divisible by 100 or equivalent
(c)	Other survey suggests that there is no point in stratifying by employment category as all categories have similar views. A sample stratified by sex would be the best. Neither the stratified sample above nor the systematic sample necessarily contain a representative proportion of each sex.	 E1 E1 E1	3	E1 stratifying by category pointless E1 representative proportion of each sex desirable E1 stratified by category not necessarily representative of sexes E1 systematic not necessarily representative of sexes *maximum 3*
	Total		**14**	

Q	Solution	Marks	Total	Comments
6(a)	280 houses	B1		
	Number houses 000 to 279	E1		OE - their total
	Select 3-digit random numbers	E1		
	Ignore repeats	E1		
	and > 279	E1	5	Consistent with their numbering
	Continue until 8 numbers obtained Select corresponding houses			
(b)	Number houses street by street, eg North St 000–062 East St 063–139 South St 140–185 West St 186–279	E1		E1 number houses street by street - may be earned in (a) but more detail required here E1 idea of systematic sampling
	Select a random number between 00 and 34.	E1		E1 choose random starting point
	Choose this house and every 35th house thereafter.	B1	3	B1 every 35th house **Maximum 3**
(c)	Cluster	B1	1	
(d)(i)	If Socrates misses a street there is a substantial probability (0.5) that John will not check any houses in this street.	E1		
(ii)	Systematic preferred	B1		
	John certain to check some houses in each street	E1	3	
(e)	No preference	B1		
	Both equally likely to check houses missed by Mary	E1	2	**SC** allow B1 for systematic because easier to carry out
	Total		**14**	

AQA_JUNE_2006_5

5(a)	number employees 0000 to 9319 select 4-digit random numbers ignore repeats and >9319 continue until 120 numbers obtained select corresponding employees	E1 E1 E1 E1	4	any valid numbering select 4-digit random numbers ignore repeats and >9319 (must be consistent in numbering) continue until 120 numbers obtained
(b)(i)	from each of the 4 chosen councils select a random sample of 30 employees	E1 E1	2	select a sample from each of the 4 councils of size 30
(b)(ii)	employees to be interviewed would be geographically localised / easier / cheaper	E2,1	2	reason – easier/cheaper without further explanation gets E1
(c)(i)	council / age / sex / length of service	B1B1	2	any sensible suggestion; B1 for each
(c)(ii)	More representative of population	E1	1	more representative allow all have equal chance
	Total		**11**	

AQA_JUNE_2008_5

5(a)(i)	Cluster	B1		
(ii)	Select 2-digit random numbers Ignore 00 and > 72 Ignore repeats Continue until 7 numbers obtained and choose passengers sitting in corresponding seats	E1 E1 E1 E1	5	If renumbered 00 to 71, max E1 E0 E1 E1
(b)(i)	Stratified	B1		Stratified / stratified random
(ii)	18:8 is ratio of number of seated standard class passengers (432) to seated first-class passengers (192) 18 + 8 = 26 in range of likely number of interviews	E1 E1 E1	4	Ratio of standard to first-class Total in right range Numerical support for ratio or demonstration that 18 and 8 is only possibility giving total in desired range
(c)(i)	No Passengers in seats numbered 49–72 have no chance of being selected	B1 E1		
(ii)	Yes, all have a chance of 3/48 of being selected	B1	3	
(d)	Xavier's sample preferred First-class and standard-class passengers fairly represented in sample	B1 E1	2	
	Total		**14**	

AQA_JUNE_2010_7

7(a)	Number shops 000 to 419 Select 3-digit random numbers Ignore repeats and > 419 Select corresponding shops	E1 E1 E1 E1	4	E1 number 000 to 419 or equivalent E1 3-digit random numbers E1 ignore >419 - consistent with their numbering E1 ignore repeats
(b)(i)	(A) stratified (random) (B) equally likely (C) not all subsets possible	E1 E1 E1	3	E1 stratified E1 equally likely Allow 'yes,'disallow 'likely' E1 reason
(ii)	(A) systematic (B) equally likely (C) not all subsets possible	E1 E1 E1	3	E1 systematic E1 equally likely. E1 reason
(iii)	Shops with largest electricity consumption are likely to have the largest potential savings. Therefore sensible to audit these first	E1 E1	2	E1 shops with largest consumption selected E1 largest potential saving
	Total		**12**	