Version : 1.0 0707

ASSESSMENT and QUALIFICATIONS ALLIANCE

General Certificate of Education

Statistics 6380

SS02 Statistics unit 2

Mark Scheme

2007 examination - June series

Mark schemes are prepared by the Principal Examiner and considered, together with the relevant questions, by a panel of subject teachers. This mark scheme includes any amendments made at the standardisation meeting attended by all examiners and is the scheme which was used by them in this examination. The standardisation meeting ensures that the mark scheme covers the candidates' responses to questions and that every examiner understands and applies it in the same correct way. As preparation for the standardisation meeting each examiner analyses a number of candidates' scripts: alternative answers not already covered by the mark scheme are discussed at the meeting and legislated for. If, after this meeting, examiners encounter unusual answers which have not been discussed at the meeting they are required to refer these to the Principal Examiner.

It must be stressed that a mark scheme is a working document, in many cases further developed and expanded on the basis of candidates' reactions to a particular paper. Assumptions about future mark schemes on the basis of one year's document should be avoided; whilst the guiding principles of assessment remain constant, details will change, depending on the content of a particular examination paper.

Set and published by the Assessment and Qualifications Alliance.

The Assessment and Qualifications Alliance (AQA) is a company limited by guarantee registered in England and Wales (company number 3644723) and a registered charity (registered charity number 1073334).
Registered address: AQA, Devas Street, Manchester M15 6EX *Dr Michael Cresswell Director General*

Key to mark scheme and abbreviations used in marking

M	mark is for method
m or dM	mark is dependent on one or more M marks and is for method
A	mark is dependent on M or m marks and is for accuracy
B	mark is independent of M or m marks and is for method and accuracy
E	mark is for explanation

√or ft or F	follow through from previous incorrect result	MC	mis-copy
CAO	correct answer only	MR	mis-read
CSO	correct solution only	RA	required accuracy
AWFW	anything which falls within	FW	further work
AWRT	anything which rounds to	ISW	ignore subsequent work
ACF	any correct form	FIW	from incorrect work
AG	answer given	BOD	given benefit of doubt
SC	special case	WR	work replaced by candidate
OE	or equivalent	FB	formulae book
A2,1	2 or 1 (or 0) accuracy marks	NOS	not on scheme
–*x* EE	deduct *x* marks for each error	G	graph
NMS	no method shown	c	candidate
PI	possibly implied	sf	significant figure(s)
SCA	substantially correct approach	dp	decimal place(s)

No Method Shown

Where the question specifically requires a particular method to be used, we must usually see evidence of use of this method for any marks to be awarded. However, there are situations in some units where part marks would be appropriate, particularly when similar techniques are involved. Your Principal Examiner will alert you to these and details will be provided on the mark scheme.

Where the answer can be reasonably obtained without showing working and it is very unlikely that the correct answer can be obtained by using an incorrect method, we must award **full marks**. However, the obvious penalty to candidates showing no working is that incorrect answers, however close, earn **no marks**.

Where a question asks the candidate to state or write down a result, no method need be shown for full marks.

Where the permitted calculator has functions which reasonably allow the solution of the question directly, the correct answer without working earns **full marks**, unless it is given to less than the degree of accuracy accepted in the mark scheme, when it gains **no marks**.

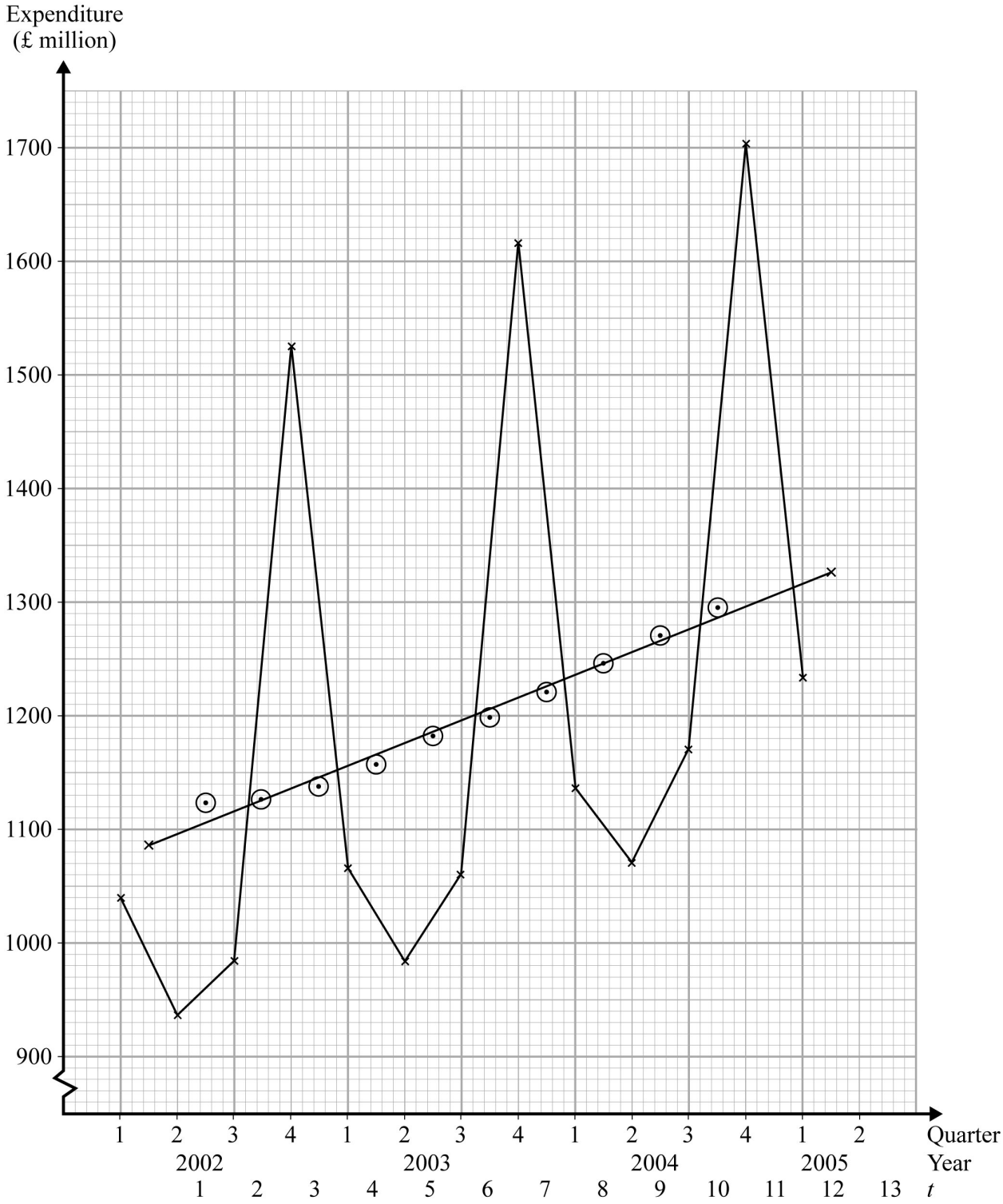
Otherwise we require evidence of a correct method for any marks to be awarded.

June 07

SS02

Q	Solution	Marks	Total	Comments
1(a)(i)	$P(3 \text{ or fewer})=0.779$	B1		$0.779\,(0.778\sim0.779)$
(a)(ii)	$P(3)=P(3 \text{ or fewer})-P(2 \text{ or fewer})$	M1		$P(3)=P(\leq3)-P(\leq2)$
	$=0.7787-0.5679$	m1		completely correct method
	$=0.209$	A1	4	$0.209(0.208\sim0.21)$
(b)	Poisson mean $5\times2.4=12$	B1		Poisson mean 5×2.4
	$P(>10)=1-P(10 \text{ or fewer})$	M1		$P(>10)=1-P(10 \text{ or fewer})$
	$=1-0.3472$			
	$=0.653$	A1	3	$0.653\,(0.652\sim0.653)$
(c)	No, customers are likely to join shortest queue i.e. not at random.	E1 E1	2	No Reason – allow not independent – couple may shop together etc.
	Total		**9**	
2(a)	$\frac{983+1059+1618+1135}{4}=1199$	M1 A1	2	method $1199\,(1198\sim1200)$
(b)	on next page	M1 A1	2	m.a. in correct position Accurate plot – by eye – allow 1 small slip
(c)	$t=0 \quad y=1086$ $t=12 \quad y=1326$ + line	M1 A1	2	method for line accurate line drawn
(d)	residuals for Q2 $-158, -193, -196$	M1		method for residual – allow from graph – ignore sign – their line
	mean = –179	m1		method for seasonal effect – ignore sign – allow omission of Q2,2002
		A1	3	$-179(-170\sim-200)$ **2 maximum if answered in (e)**
(e)	$1086+12.5\times19.96-179$	M1		method for trend – allow them from graph – their line
	$=1335.5-179$	M1		method for including their negative seasonal effect – their trend
	$=1156.5$	B1		$(1130\sim1170)$ allow 1100 or 1200
	forecast £1160 million	B1√	4	2 or 3 sf and £m **Allow 3 maximum if method is not clear or based on Q2 results only**
(f)	this is a poor forecast but no purely numerical method could have predicted Q2,2005 would be less than Q2,2004	E1 E1	2	poor forecast / ineffective method no numerical method could have forecast this result / extrapolation is inherently unreliable
	Total		**15**	

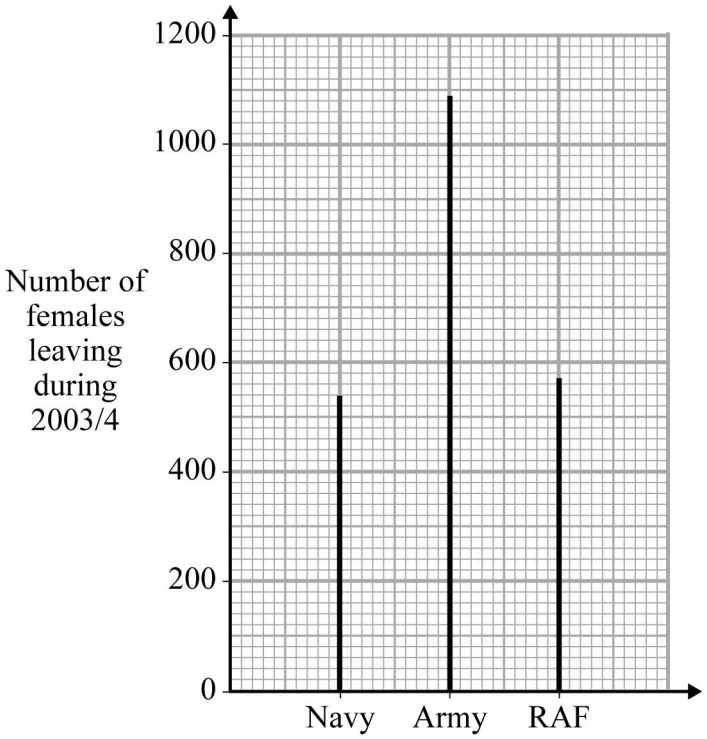
SS02 (cont)


Expenditure
(£ million)
1700
1600
1500
1400
1300
1200
1100
1000
900
1 2 3 4 1 2 3 4 1 2 3 4 1 2
Quarter
2002 2003 2004 2005
Year
1 2 3 4 5 6 7 8 9 10 11 12 13
t

SS02 (cont)

Q	Solution	Marks	Total	Comments
3(a)	$\mathrm{E}(X) = 0\times0.32 + 1\times0.25 + 2\times0.19 + 3\times0.12 + 4\times0.09 + 5\times0.03$	M1		method for E(X)
	= 1.5	A1		1.5 CAO
	$\mathrm{E}(X^2) = 0^2\times0.32 + 1^2\times0.25 + 2^2\times0.19 + 3^2\times0.12 + 4^2\times0.09 + 5^2\times0.03$	M1		method for E(X^2) – may be implied
	= 4.28			
	$\mathrm{Var}(X) = 4.28 - 1.5^2 = 2.03$	m1		method for s.d.; allow for variance = 2.03
	s.d. $= \sqrt{2.03}$			
	= 1.42	A1	5	$1.42(1.41\sim1.43)$
(b)(i)	s.d. $= \sqrt{2.2}$	M1		method
	=1.48	A1	2	$1.48(1.48\sim1.49)$
(b)(ii)	more houses in Cheadleville are advertised in the Clarion than in the Sentinel. The week to week variability is similar	E1√		Clarion higher average
		E1	2	variability similar
(c)	choose Clarion – since more houses in Cheadleville advertised on average	B1√		Clarion
		B1	2	higher mean
	Total		**11**	
4(a)	15320	B1	1	15320 or 15300
(b)	890 – 580 = 310	M1		method
		A1	2	310 CAO
(c)	1998/9 to 2003/4	M1		method – allow small slip
		A1	2	1998/9 to 2003/4 CAO
(d)	on next page	M1		method – allow horizontal – allow bars instead of lines but not if joined – disallow broken scale
		B1		axes labelled – generous
		A1	3	accurate plot by eye
(e)(i)	$\frac{11950}{32130}\times100 = 37.2\%$	M1		method for ratio
		A1		$37.2(37\sim37.3)$
(e)(ii)	more leaving than joining – in the long run this would lead to no one being left to leave	E1	3	explanation
	Total		**11**	

SS02 (cont)


1200
1000
800
600
400
200
0
Number of females leaving during 2003/4
Navy
Army
RAF

SS02 (cont)

Q	Solution	Marks	Total	Comments
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**	

SS02 (cont)

Q	Solution	Marks	Total	Comments
6(a)(i)	$H_0 : \mu = 41 \quad H_1 : \mu > 41$	B1 B1		correct hypothesis - generous both hypotheses correct – requires population or μ
	$\bar{x}$=52.03	B1		$52.03 (52 \sim 52.1)$
	$z = \frac{52.03-41}{\frac{8.5}{\sqrt{10}}} = 4.10$	M1 m1 A1		use of $\frac{8.5}{\sqrt{10}}$ correct method for z $4.10 (4.10 \sim 4.11)$
	c.v. 2.3263	B1		$2.3263 (2.32 \sim 2.33)$ – ignore sign
	reject H_0: significant evidence that mean speed exceeds 41 mph	A1√	8	conclusion in context AG – must be compared with upper tail of z
	non-standardised c.v. $41+2.3263\times\left(\frac{8.5}{\sqrt{10}}\right)=47.25$ compare with 52.03 **confidence interval** $52.03\pm2.3263\times\frac{8.5}{\sqrt{10}}$ 45.78~58.28 compare 45.78 with 41 ***p*-value** compare 0.0000204 with 0.01			
(a)(ii)	not a random sample – it contains only drivers prosecuted for speeding, who will be the fastest	E1 E1	2	not random reason

SS02 (cont)

Q	Solution	Marks	Total	Comments
(b)(i)	$H_0 : \mu = 30 \quad H_1 : \mu > 30$	B1		both hypotheses
	$z = \frac{31.6-30}{\frac{6.9}{\sqrt{120}}} = 2.54$	M1 A1		method for z – ignore sign $2.54(2.53\sim2.55)$
	c.v. = 1.6449	B1		$1.6449(1.64\sim1.65)$ – ignore sign
	reject H_0: significant evidence that mean speed exceeds 30 mph	A1√	5	ft conclusion in context – must be compared with upper tail of z
	non-standardised c.v. $30+1.6449\times\frac{6.9}{\sqrt{120}}=31.04$ compare with 31.6 **confidence interval** $31.6\pm1.6449\times\frac{6.9}{\sqrt{120}}$ 30.56~32.64 compare 30.56 with 30 ***p*-value** compare 0.00554 with 0.05			
(b)(ii)	mean speed above 30 – indicates most cars probably above limit – although distribution probably skew so most may be below limit. since 31.6 significantly higher than 30 it will certainly be significantly lower than 41 so average speed has been reduced	E3,2,1	3	mean above limit/most speeding/ distribution skew/average speed reduced any sensible comments; E1 for each upto maximum 3
	Total		**18**	
	TOTAL		**75**	