Version 1.0: 0606

General Certificate of Education

Statistics 6380

SS05 Statistics 5

Mark Scheme

2006 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.

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.

SS05

Q	Solution	Marks	Total	Comments
1(a)	$s^2 = 0.3955$	B2	2	AWFW 0.395 to 0.396
(b)	Assuming the weights of truffles are normally distributed:	B1		Must indicate that population is normally distributed
	$\nu = 12 - 1 = 11$	B1		
	CVs for χ^2 are 3.816, 21.920	B1		Both; accept 3.82, 21.9
	confidence limits for σ^2 are $\frac{11\times 0.3955}{21.920}$, $\frac{11\times 0.3955}{3.816}$	M1 A1√		√ on s^2 and χ^2 values
	giving confidence interval (0.198, 1.14)	A1	6	(0.198 to 0.199, AWRT)
(c)	sd of 0.4 $\Rightarrow$ variance of 0.16 which is below CI so statement is unlikely to be true.	E1√ B1√	2	√ on CI √ on CI B1 if CI for standard deviation found but incorrectly used.
	Total		**10**	
2(a)	X = time in minutes for bus 6 journey Y = time in minutes for bus 23 journey			
	H_0: $\mu_X = \mu_Y$	B1		Compares population means
	H_1: $\mu_X \neq \mu_Y$	B1		= and ≠
	CVs of z are ±1.96	B1		1.96 accepted as implying 2-tail
	test statistic = $\frac{13.2 - 14.6}{\sqrt{\frac{1.8^2}{15} + \frac{1.8^2}{10}}}$	M1 A1		Subtraction either way if used consistently
	$= -1.905$	A1		AWFW –1.91 to –1.90
	Not enough evidence to suggest a difference in mean journey times.	A1√	7	√ on CV and test statistic Only if 2-tail test used
(b)	Random samples from population of journey times or samples are independent.	B1	1	Accept valid comment in context implying samples are representative
(c)	Concluding there is no difference in mean journey times when there is a difference.	B1	1	
	Total		**9**	

SS05 (cont)

Q	Solution	Marks	Total	Comments
3(a)(i)	$\text{f}(x)=\begin{cases}0.2 & 97.5\le x\le 102.5\\ 0 & \text{otherwise}\end{cases}$	B1	1	Mark given for finding 0.2 without formal expression
(ii)	$\text{sd}=\sqrt{\dfrac{(102.5-97.5)^2}{12}}$	M1		
	$=\dfrac{5}{\sqrt{12}}=1.44$	A1	2	AWRT
(b)	$\text{P}(98\le X\le 101)=\dfrac{101-98}{102.5-97.5}$	M1		
	$=\dfrac{3}{5}=0.6$	A1	2	CAO
(c)	$\bar{X}\sim N\left(100,\dfrac{25}{12\times 50}\right)$	B1 B1		Normal distribution Mean = 100; Variance from (a)(ii) divided by 50
	$=N\left(100,\dfrac{1}{24}\right)$			
	Large sample so central limit theorem applies.	E1	3	Mark given for mean of large sample or reference to CLT
	Total		**8**	

SS05 (cont)

Q	Solution	Marks	Total	Comments
4(a)	H_0: Number of calls, X, ~ Poisson(3)	B1		Must specify which Poisson
	H_1: not H_0			In table: accept probabilities by formula or tables to at least 3 dp; E values to at least 2 dp.
	x O $P(X=x)$ E $\frac{(O-E)^2}{E}$			
	0 1 0.0498 3.984 } 2.206	M1		Poisson probabilities
	1 9 0.1493 11.944 }	M1		E = Probability × 80
	2 23 0.2241 17.928 1.435	M1		Combining first two groups
	3 19 0.2240 17.920 0.065			(O =10; E = 15.92 to 15.93)
	4 17 0.1681 13.448 0.938	m1		Last group ≥ 6; dependent on first M1
	5 6 0.1008 8.064 0.528	A2		7 Es AWRT 1 dp (A1 if 5 or 6); lose one A mark for too much rounding
	≥ 6 5 0.0839 6.712 0.437			
	80 1 80 5.609			
	$\chi^2 = \sum \frac{(O-E)^2}{E} = 5.609$	M1 A1√		Use of formula AWFW 5.60 to 5.71; √ only on minor calculation error
	DF $\nu = (7-1)-1 = 5$	B1		No. of categories used –1
	CV of χ^2 = 9.236	B1√		√ on ν
	Insufficient evidence at 10% level to say that Poisson (3) is not a suitable model.	A1√	12	√ on CV and tabulated total Accept Poisson with no parameter
(b)(i)	Same number of categories but one more constraint so ν = 4, giving	E1		1 less than DF in (a)
	χ^2 = 7.779	B1	2	χ^2 matching ν, seen here or in (ii)
(ii)	Total from table unchanged and is < 7.779 so conclusion is same as in (a)	B1√	1	√ on values from (a) if consistent; conclusion must be justified
(c)	Number of calls, Y, in 8 hours ~ Poisson(24) ≈ N(24, 24)	B1		
	$P(Y > 30) = 1 - \Phi\left(\frac{30.5-24}{\sqrt{24}}\right)$	M1 A1		Continuity correction
	$= 1 - \Phi(1.33)$			
	$= 0.092$	A1√	4	√ on lack of c. c. only SC B1 if try to use normal approximation, but wrong variance
	Alternatively By calculator from Poisson: $P(Y > 30) = 1 - P(Y \le 30)$ $= 1 - 0.90415 = 0.0959$	(B4)		AWRT 0.096
	Total		**19**	

SS05 (cont)

Q	Solution	Marks	Total	Comments
5(a)(i)	$\mathrm{E}(T) = \dfrac{1}{0.8}$	M1		
	$= 1.25$	A1	2	
(ii)	Mean lifetime = $1.25 \times 1000 = 1250$ hours	A1√	1	√ on E(*T*) Condone lack of units
(iii)	P(lifetime < 1000) = P($T < 1$)	B1		Using $t = 1$ or mean = 1250 Or $\lambda = 0.0008$
	$\mathrm{P}(T \le t) = 1 - \mathrm{e}^{-0.8t}$			
	$\mathrm{P}(T \le 1) = 1 - \mathrm{e}^{-0.8}$	M1		Attempt to use F(*x*) or F(*t*) or integration
	$= 0.551$	A1	3	0.550 to 0.551
(b)(i)	Exponential distribution has 'no memory' so required probability is P($T > 0.5$)	M1		Use of this property; P($T > 1$) or $\mathrm{P}(X > 500)$ seen
	$= 1 - \left(1 - \mathrm{e}^{-0.4}\right)$	m1		1–F (0.5) or F(500) attempted
	$= 0.670$	A1	3	0.670 to 0.671
(ii)	P(at least one fails within 500 hours) = 1 – P(all last more than 500 hours)	M1		Or combining probs. For 1,2 and 3 failures
	$= 1 - \left(\mathrm{e}^{-0.4}\right)^3$			
	$= 1 - (0.670)^3$	m1		Or use of multiplication and addition laws
	$= 0.699$	A1√	3	AWRT 0.70; √ on sensible probability
	Total		**12**	

SS05 (cont)

Q	Solution	Marks	Total	Comments
6(a)	H0: $\sigma_X{}^2 = \sigma_Y{}^2$ H1: $\sigma_X{}^2 \neq \sigma_Y{}^2$	B1		Or H0: $\sigma_X = \sigma_Y$ H1: $\sigma_X \neq \sigma_Y$; (both)
	DF: $\nu_1 = 10$, $\nu_2 = 8$	B1		
	CV of F is 3.347	B1		Both B1 if dof reversed with matching CV
	$\frac{s_X{}^2}{s_Y{}^2} = \frac{3.798}{2.925}$	M1		
	= 1.30	A1		AWRT
	< CV so reasonable to believe that $\sigma_X{}^2 = \sigma_Y{}^2$	E1	6	Justifies conclusion; accept appropriate diagram
(b)	H0: $\mu_X - \mu_Y = 10$ H1: $\mu_X - \mu_Y < 10$	B2,1		B1 For saying difference of means = 10; <10 without saying which way subtracted or for X and Y reversed
	Pooled estimate of variance $= \frac{(10\times 3.798)+(8\times 2.925)}{18}$	M1		
	= 3.41	A1		
	DF $\nu = 18$	B1		
	CV of $t = -2.552$	B1		Accept + 2.552 here
	test statistic = $\frac{(36.8-29.1)-10}{\sqrt{3.41\left(\frac{1}{11}+\frac{1}{9}\right)}}$	M1 A1 A1√		Use of correct formula For –10 Correct values substituted; √ on pooled estimate
	= –2.77	A1√		√ on pooled estimate
	< –2.552 so supports Stephen's claim that the difference in mean fat contents is less than 10 grams.	A1√	11	√ on test statistic and CV
	Total		**17**	
	TOTAL		**75**	