GCE Examinations

Advanced Subsidiary / Advanced Level

Statistics
Module S3

Paper B

MARKING GUIDE

This guide is intended to be as helpful as possible to teachers by providing concise solutions and indicating how marks should be awarded. There are obviously alternative methods that would also gain full marks.

Method marks (M) are awarded for knowing and using a method.

Accuracy marks (A) can only be awarded when a correct method has been used.

(B) marks are independent of method marks.

Written by Shaun Armstrong & Chris Huffer

S3 Paper B – Marking Guide

1. *(a)* divide poulation into distinct groups
sample sizes from each group determined by proportions in population
sample any member of group until quota is filled — B3

(b) e.g. non-random sample within strata so may be biased — B1

(c) e.g. survey on political attitudes according to age group
too time-consuming / impractical to random sample within strata — B2 **(6)**

---

2. exp. freq. 0-45 = $\frac{45}{360} \times 96 = 12$ etc. giving exp. freqs. 12, 12, 24, 24, 12, 12 — M1 A1

H_0 : continuous uniform distribution is a suitable model
H_1 : continuous uniform distribution is not a suitable model — B1

O	E	$(O-E)$	$\frac{(O-E)^2}{E}$
18	12	6	3
19	12	7	4.0833
15	24	$^-9$	3.375
20	24	$^-4$	0.6667
9	12	$^-3$	0.75
15	12	3	0.75

$\therefore \Sigma \frac{(O-E)^2}{E} = 12.625$ — M1 A2

$\nu = 6 - 1 = 5$, $\chi^2_{crit}(5\%) = 11.070$ — M1 A1

$12.625 > 11.070 \;\therefore$ reject H_0

continuous uniform distribution is not a suitable model — A1 **(9)**

---

3. *(a)* mean = $\frac{(46\times60)+15}{20} = 138.75$ — M1

C.I. $\bar{x} \pm 1.6449\frac{\sigma}{\sqrt{n}} = 138.75 \pm 1.6449.\frac{23}{\sqrt{20}}$ — M1 A1

giving (130.3, 147.2) — A2

(b) width = $2 \times 1.6449 \times \frac{23}{\sqrt{n}} \;\therefore\; 2 \times 1.6449 \times \frac{23}{\sqrt{n}} < 10$ — M1 A1

$\therefore \sqrt{n} > 7.56654$ — A1

giving $n > 57.25$ so min. value of $n = 58$ — M1 A1

(c) e.g. she might buy big-budget movies with longer credits — B1 **(11)**

---

4. expected freq. 8am-6pm/minor = $\frac{108\times56}{148} = 40.86$

6pm-2am/minor = $\frac{108\times71}{148} = 51.81$ — M1 A2

giving expected freqs

40.86	15.14
51.81	19.19
15.33	5.67

— A1

H_0 : proportion of serious injuries independent of time
H_1 : proportion of serious injuries varies with time — B1

O	E	$(O-E)$	$\frac{(O-E)^2}{E}$
45	40.86	4.14	0.4195
11	15.14	$^-4.14$	1.1321
49	51.81	$^-2.81$	0.1524
22	19.19	2.81	0.4115
14	15.33	$^-1.33$	0.1154
7	5.67	1.33	0.3120

$\therefore \Sigma \frac{(O-E)^2}{E} = 2.543$ — M1 A2

$\nu = 2$, $\chi^2_{crit}(5\%) = 5.991$ — M1 A1

$2.543 < 5.991 \;\therefore$ not significant

there is no evidence of prop'n of serious injuries varying with time — A1 **(11)**

---

5. *(a)*

bottle	A	B	C	D	E	F	G	H	I	J
enth. rank	4	7	2	1	8	6	5	10	9	3
price rank	1	6	2	3	10	7	9	4	8	5
d^2	9	1	0	4	4	1	16	36	1	4

$\Sigma d^2 = 76$ — M2 A2

$r_s = 1 - \frac{6\times76}{10\times99} = 0.5394$ — M1 A1

(b) $H_0 : \rho = 0 \quad H_1 : \rho > 0$ — B1

$n = 10$, 5% level $\therefore$ C.R. is $r_s > 0.5636$ — M1 A1

$0.5394 < 0.5636$ $\therefore$ not significant

there is no evidence of positive correlation — A1

(c) share ranks, both 6.5, use pmcc — B2 **(12)**

---

6. *(a)* $\hat{\mu} = \overline{V} = \frac{10367}{80} = 129.6$ cm — M1 A1

$\hat{\sigma}^2 = s^2 = \frac{80}{79}\left(\frac{1350314}{80} - 129.5875^2\right) = 87.09$ — M2 A1

(b) $H_0 : \mu_V = \mu_M \quad H_1 : \mu_V \neq \mu_M$ — B1

1% level $\therefore$ C.R. is $z < {}^-2.5758$ or $z > 2.5758$ — B1

test statistic $= \frac{129.6-130.5}{\sqrt{\frac{87.09}{80}+\frac{96.24}{280}}} = {}^-0.7520$ — M2 A2

not in C.R. do not reject H_0 — M1

no evidence of difference in mean heights — A1 **(13)**

---

7. *(a)* let X = time to mark P1 paper

let $A = X_1 - X_2$ $\therefore$ $A \sim N(0, 2\times17^2) = \sim N(0, 578)$ — M1 A1

$P({}^-5 < A < 5) = P\left(\frac{-5-0}{\sqrt{578}} < Z < \frac{5-0}{\sqrt{578}}\right)$ — M1 A1

$= P({}^-0.21 < Z < 0.21) = 0.5832 - (1 - 0.5832) = 0.166$ — M1 A1

(b) let M = time to mark M1 paper, let S = time to mark S1 paper

let $T = M_1 + \ldots + M_{45} + S_1 + \ldots + S_{80}$

$\therefore$ $T \sim N(45\times314 + 80\times284, 45\times42^2 + 80\times29^2) = \sim N(36\,850, 146\,660)$ — M2 A2

P(time < 10 hours) $= P(T < 36\,000) = P\left(Z < \frac{36000-36850}{\sqrt{146660}}\right)$ — M1

$= P(Z < {}^-2.22) = 1 - 0.9868 = 0.0132$ — M1 A1 **(13)**

Total **(75)**

Performance Record – S3 Paper B

Question no.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	Total
Topic(s)	sampling	goodness of fit, cont. unif.	confidence interval	conting. table	Spearman's, hyp. test	unbiased estimates, diff. of means hyp. test	linear comb. of Normal r.v.	
Marks	6	9	11	11	12	13	13	75
Student								