GCE Examinations

Advanced Subsidiary / Advanced Level

Statistics

Module S3

Paper D

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 D – Marking Guide

1. *(a)* list volunteers
from random pt in table look at 2-digit nos until get one from 01 to 12
take this one from list and then every 12th person on list — B3

(b) e.g. advantage – quicker
disadvantage – not random unless list is, so may introduce bias — B2 **(5)**

---

2. *(a)* $\bar{x} = \frac{1419}{30} = 47.3$ — M1

C.I. is $\bar{x} \pm 1.96 \frac{\sigma}{\sqrt{n}} = 47.3 \pm 1.96 . \frac{5}{\sqrt{30}}$ — M1 A1

giving (45.51, 49.09) — A2

(b) $\frac{19}{20}$ — B1

(c) it either does or doesn't include true mean ∴ probability is 0 or 1 — B1 **(7)**

---

3. *(a)*

candidate	*A*	*B*	*C*	*D*	*E*	*F*
exp. rank	3	1	4	6	5	2
new rank	6	3	2	4	5	1
d^2	9	4	4	4	0	1

$\Sigma d^2 = 22$ — M2 A1

$r_s = 1 - \frac{6 \times 22}{6 \times 35} = 0.3714$ — M1 A1

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

$n = 6$, 5% level ∴ C.R. is $r_s > 0.8286$ — M1 A1

$0.3714 < 0.8286$ ∴ not significant
there is no evidence of positive correlation — A1

(c) e.g. needs training as assessment not in line with experienced manager — B1 **(10)**

---

4. *(a)* $\hat{\mu} = \bar{t} = \frac{1039}{30} = 34.6$ — M1 A1

$\hat{\sigma}^2 = s^2 = \frac{30}{29} \left(\frac{65393}{30} - 34.633^2 \right) = 1014.1$ — M1 A1

(b) $\frac{\Sigma x}{20} = 32.0 \quad \therefore \Sigma x = 640 \quad \hat{\mu}$ for combined sample $= \frac{1039+640}{50} = 33.6$ — M1 A1

$963.4 = \frac{20}{19} \left(\frac{\Sigma x^2}{20} - 32.0^2 \right)$ giving $\Sigma x^2 = 38784.6$ — M1 A1

$\hat{\sigma}^2$ for combined sample $= \frac{50}{49} \left(\frac{65393+38784.6}{50} - 33.58^2 \right) = 975.4$ — M1 A1 **(10)**

---

5. *(a)* let W = weight of egg
let $A = W_1 - W_2 \quad \therefore A \sim N(0, 2 \times 3.9^2) = \sim N(0, 30.42)$ — M1 A1

require $2 \times P(A > 4) = 2 \times P\left(Z > \frac{4-0}{\sqrt{30.42}}\right)$ — M1

$= 2 \times P(Z > 0.73) = 2 \times (1 - 0.7673) = 0.465$ — M1 A1

(b) let T = total weight of box and eggs
$\therefore T \sim N(28 + 6 \times 55, 1.2^2 + 6 \times 3.9^2) = \sim N(358, 92.7)$ — M1 A2

$P(T < 350) = P\left(Z < \frac{350-358}{\sqrt{92.7}}\right)$ — M1

$= P(Z < -0.83) = 1 - 0.7967 = 0.2033$ — M1 A1 **(11)**

---

6. *(a)*

	accident	no accident	
< 25 yrs	104	216	320
≥ 25 yrs	16	64	80
	120	280	400

M1 A1

(b) (i) expected freq. < 25/accident = $\frac{120\times320}{400}$ = 96 M1 A1

giving expected freqs 96 224
24 56 A1

H_0 : no assoc'n between age pass test and accident in next 2 yrs
H_1 : there is assoc'n between age pass test and acc in next 2 yrs B1

O	E	$(O-E)$	$\frac{(O-E)^2}{E}$
104	96	8	0.6667
216	224	−8	0.2857
16	24	−8	2.6667
64	56	8	1.1429

$\therefore \Sigma\frac{(O-E)^2}{E} = 4.762$ M1 A2

$\nu = 1, \chi^2_{\text{crit}}(5\%) = 3.841$ M1 A1

$4.762 > 3.841$ $\therefore$ significant
evidence of assoc'n between age pass test and acc in next 2 yrs A1

(ii) using totals, which must agree, once know one value
can calculate all others B1

(c) higher proportion of accidents in < 25 led to significant result
extra data increases this difference so still significant B2 **(15)**

---

7. *(a)* let X = length of adult male feet

$P(21.5 < X < 24.5) = P(\frac{21.5-22.4}{2.8} < Z < \frac{24.5-22.4}{2.8})$ M1

$= P(-0.32 < Z < 0.75) = 0.7734 - (1 - 0.6255) = 0.3989$ M1 A1

exp. freq. = $0.3989 \times 200 = 79.78$ A1

$P(24.5 < X < 27.5) = P(0.75 < Z < \frac{27.5-22.4}{2.8})$

$= P(0.75 < Z < 1.82) = 0.9656 - 0.7734 = 0.1922$ M1

exp. freq. = $0.1922 \times 200 = 38.44$ A1

exp. freq. for > 27.5 = 200 − total of others = 6.88 A1

(b) H_0 : $N(22.4, 2.8^2)$ is a suitable model
H_1 : $N(22.4, 2.8^2)$ is not a suitable model B1

O	E	$(O-E)$	$\frac{(O-E)^2}{E}$
24	16.46	7.54	3.4539
48	58.44	−10.44	1.8651
69	79.78	−10.78	1.4566
41	38.44	2.56	0.1705
18	6.88	11.12	17.9730

$\therefore \Sigma\frac{(O-E)^2}{E} = 24.919$ M1 A2

$\nu = 5 - 1 = 4, \chi^2_{\text{crit}}(10\%) = 7.779$ M1 A1

$24.919 > 7.779$ $\therefore$ reject H_0
$N(22.4, 2.8^2)$ is not a suitable model A1

(c) use data to estimate mean and std. dev.
combine any cells with exp. freqs. < 5 and repeat calculation
ν = no of cells after combining − 3 as parameters have been estimated B3 **(17)**

Total **(75)**

Performance Record – S3 Paper D

Question no.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	Total
Topic(s)	sampling	confidence interval	Spearman's, hyp. test	unbiased estimates	linear comb. of Normal r.v.	conting. table	goodness of fit, Normal	
Marks	5	7	10	10	11	15	17	75
Student								