



*Rewarding Learning*

**ADVANCED**  
**General Certificate of Education**  
**2015**

---

**Mathematics**

Assessment Unit S4

*assessing*

Module S2: Statistics 2

[AMS41]

**TUESDAY 16 JUNE, AFTERNOON**

---

**MARK**  
**SCHEME**

## GCE ADVANCED/ADVANCED SUBSIDIARY (AS) MATHEMATICS

### Introduction

The mark scheme normally provides the most popular solution to each question. Other solutions given by candidates are evaluated and credit given as appropriate; these alternative methods are not usually illustrated in the published mark scheme.

The marks awarded for each question are shown in the right-hand column and they are prefixed by the letters **M**, **W** and **MW** as appropriate. The key to the mark scheme is given below:

**M** indicates marks for correct method.

**W** indicates marks for working.

**MW** indicates marks for combined method and working.

The solution to a question gains marks for correct method and marks for an accurate working based on this method. Where the method is not correct no marks can be given.

A later part of a question may require a candidate to use an answer obtained from an earlier part of the same question. A candidate who gets the wrong answer to the earlier part and goes on to the later part is naturally unaware that the wrong data is being used and is actually undertaking the solution of a parallel problem from the point at which the error occurred. If such a candidate continues to apply correct method, then the candidate's individual working must be followed through from the error. If no further errors are made, then the candidate is penalised only for the initial error. Solutions containing two or more working or transcription errors are treated in the same way. This process is usually referred to as "follow-through marking" and allows a candidate to gain credit for that part of a solution which follows a working or transcription error.

### Positive marking:

It is our intention to reward candidates for any demonstration of relevant knowledge, skills or understanding. For this reason we adopt a policy of **following through** their answers, that is, having penalised a candidate for an error, we mark the succeeding parts of the question using the candidate's value or answers and award marks accordingly.

Some common examples of this occur in the following cases:

- (a) a numerical error in one entry in a table of values might lead to several answers being incorrect, but these might not be essentially separate errors;
- (b) readings taken from candidates' inaccurate graphs may not agree with the answers expected but might be consistent with the graphs drawn.

When the candidate misreads a question in such a way as to make the question easier only a proportion of the marks will be available (based on the professional judgement of the examining team).

		AVAILABLE MARKS
1	(a) (i) no effect	MW1
	(ii) no effect	MW1
1	(b) (i) $r = \frac{\sum xy - \frac{\sum x \cdot \sum y}{n}}{\sqrt{\sum x^2 - \frac{(\sum x)^2}{n}} \sqrt{\sum y^2 - \frac{(\sum y)^2}{n}}}$	M1
	$= \frac{186.3 - \frac{(74.6)(24.7)}{10}}{\sqrt{560 - \frac{(74.6)^2}{10}} \sqrt{63.17 - \frac{(24.7)^2}{10}}}$	MW2
	$= \frac{2.038}{\sqrt{3.484} \sqrt{2.161}}$	
	$= 0.743 \quad (3 \text{ s.f.})$	W1
	(ii) Moderate positive correlation, i.e. evidence that there is a possible relationship between $x$ and $y$ .	MW1
		7
2	(i) $x$ – number of thefts $\sum x = 30$	
	$\sum xf = 130 \quad \sum x^2 f = 700$	
	$\bar{x} = \frac{130}{30} = \frac{13}{3}$	M1 W1
	$\hat{\sigma}^2 = \frac{\sum x^2 f - n\bar{x}^2}{n-1} = \frac{700 - 30 \cdot \left(\frac{13}{3}\right)^2}{29} = 4.71 \text{ (3 s.f.)}$	M1 W1
2	(ii) New sample $n = 50$	
	2nd sample $\sum y = 20 \times 5 = 100; \quad 5 = \frac{\sum y^2 - 20 \times 5^2}{19}$	MW1
	$\sum y^2 = 595$	MW1
	New sample $\bar{x} = \frac{130 + 100}{50} = 4.60$	MW1 W1
	$\hat{\sigma}^2 = \frac{1295 - 50 \times (4.60)^2}{49} = 4.84 \text{ (3 s.f.)}$	MW1 W1
		10

- 3 Let  $L$  be the time to clean limousine  
 Let  $V_1$  be the time to clean vintage car 1  
 Let  $V_2$  be the time to clean vintage car 2

$$L \approx N(25, 36)$$

$$V_1 \approx N(15, 16)$$

$$V_2 \approx N(15, 16)$$

Event is  $D = L - V_1 - V_2$

Prob ( $L > V_1 + V_2$ ) =  $P(D > 0)$

$$\mu_D = 25 - 30 = -5$$

$$\sigma_D^2 = 36 + 16 + 16 = 68$$

$$\text{Prob}(D > 0) = P\left[Z > \frac{0 - (-5)}{\sqrt{68}}\right]$$

$$= P[Z > 0.6063] = 1 - 0.7276$$

$$= 0.272 \quad (3 \text{ s.f.})$$

MW1

MW1

W1

M1 W1

M1

MW1

W1

8

4 (i)  $y = a + bx$        $b = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}}$

$$= \frac{458.1 - \frac{66 \times 52.26}{8}}{639 - \frac{66^2}{8}}$$

$$= \frac{26.955}{94.5} = 0.285 \quad (3 \text{ s.f.})$$

M1

W1

$$a = \bar{y} - b \bar{x}$$

M1

$$= \frac{52.26}{8} - 0.2852 \times \frac{66}{8}$$

$$= 4.1796 = 4.18 \quad (3 \text{ s.f.}) \quad y = 4.18 + 0.285x$$

W1 W1

- (ii)  $a = 4.18$  – yield of hay per hectare if no fertilisers applied (not reliable) MW1  
 $b = 0.285$  – amount of hay for each extra tonne applied MW1

- (iii)  $x = 7$      $y = 6.18$  tonnes of hay    (3 s.f.) MW1 W1

- (iv) Last three terms indicate same yield, i.e. no increase in hay for increase in fertiliser. Data not linear in nature. MW1

10

AVAILABLE  
MARKS

			AVAILABLE MARKS	
5	$H_0: \mu = 18.1$	M1		
	$H_1: \mu > 18.1$	M1		
	One-tailed @ 5% $Z_{\text{crit}} = 1.645$	MW1		
	$Z = \frac{\bar{x} - \mu}{\frac{\hat{\sigma}}{\sqrt{n}}} = \frac{18.5 - 18.1}{\frac{\sqrt{2.54}}{\sqrt{50}}}$	M1 W1		
	$= \frac{0.4}{0.2254} = 1.775$	W1		
	Since $1.775 > 1.645$ we reject $H_0$ and there is significant evidence to suggest that the joining age has increased.	MW1		
				7
6	(i) Let $T = X_1 + X_2 + \dots + X_{10} + Y$ $X$ – box $Y$ – carton	M1		
	$\bar{T} = 7000$ grams	W1		
	$\text{Var } \bar{T} = 10 \times 144 + 30^2 = 2340$	M1		
	Standard deviation = 48.4 grams    (3 s.f.)	MW1		
	(ii) Prob ( $6900 < \bar{T} < 7100$ ) $Z = \frac{X - \mu}{\sigma}$	M1		
	$= \text{Prob}(-2.067 < Z < 2.067)$ Prob ( $\bar{T} < 7100$ ) = 0.9806	W1 W1		
	$= 0.9806 - (1 - 0.9806)$			
	$= 0.961$ (3 s.f.)	MW1	8	
7	Paired test, i.e. differences	M1		
	$\sum d = 19$ $\bar{x}_D = 2.375$	W1		
	$\sum d^2 = 341$ $s_d^2 = \frac{341 - 8(2.375)^2}{7}$	W1		
	$= 42.27$			
	$\therefore s_d = 6.50$ (3 s.f.)	MW1		
	Null Hypothesis $\mu_d = 0$	M1		
	Alternative Hypothesis $\mu_d \neq 0$ (two-tailed)	M1		
	Degrees of freedom = 7	M1		
	$t_7$ (5%) = 2.365	MW1		
	$t_7 = \frac{2.375}{\frac{6.50}{\sqrt{8}}} = 1.03$	M1 W1		
	$1.03 < 2.365$	MW1		
	There is insufficient evidence to reject the null hypothesis that the papers were of the same level of difficulty.	MW1	12	

- 8 (a) The probability that the interval contains the mean  $\mu$  is 0.95  
(It is **not** the probability that  $\mu$  lies in the interval is 0.95  
 $\mu$  is fixed, the interval is not.)

MW1

(b) (i)  $\bar{x} = \frac{107.84 + 92.16}{2} = 100$

MW1

$$100 + \frac{1.96}{\sqrt{100}} \hat{\sigma} = 107.84 \quad \therefore \quad \hat{\sigma} = 40$$

M1 W1

or

$$100 - \frac{1.96}{\sqrt{100}} \hat{\sigma} = 92.16$$

- (ii) New confidence interval would be (95, 105)

MW1

(iii) Prob ( $95 < X < 105$ )                      Prob ( $X < 105$ ) =  $P\left(Z < \frac{5}{\frac{40}{10}}\right)$

M1

$$= P(Z < 1.25)$$

$$= 0.8944$$

MW1

$$\text{Prob} (95 < X < 105) = 0.8944 - (1 - 0.8944)$$

M1

$$= 0.7888$$

$$\text{Confidence level} = 78.9\%$$

W1

(iv)  $100 + 1.96 \times \frac{40}{\sqrt{n}} = 105$

M1 W1

$$\sqrt{n} = 15.68 \quad n = 245.86$$

W1

$$\text{Smallest sample} = 246$$

W1

13

**Total**

**75**

AVAILABLE  
MARKS