

# Topic Y3 Correlation regression & chi-squared tests (Post-TT B) [42] MS

1.

(i)	$(48 \times 72/150)$ or $(48/150)(72/150) \times 150$	M1 A1	2	Multiply and divide relevant values All correct
(ii)	No, no expected value less than 5		B1	1
(iii)	$H_0$ : Volume and day are independent $(H_1$ : Volume and day are not independent) Critical value for 4 df=13.28 Test statistic > 13.28, reject $H_0$ Accept that volume and day are not independent	B1 B1 M1 A1	B1	Attributes specified 4
(iv)	Choose Friday Highest volume	B1	B1	2 Not reference to E values

2.

(i)	$\Sigma x = 1366$ $\Sigma y = 17.6$ $\Sigma x^2 = 374460$ $\Sigma y^2 = 62.82$ $\Sigma xy = 4784.8$  $S_{xx} = 374460 - \frac{1366^2}{5}$ or 1268.8 $S_{yy} = 62.82 - \frac{17.6^2}{5}$ or 0.868 $S_{xy} = 4784.8 - \frac{1366 \times 17.6}{5}$ or -23.52  $r = \frac{-23.52}{\sqrt{1268.8 \times 0.868}}$ or $\frac{-23.52}{33.186...}$ oe = -0.709 (3 sfs)	B1  M1  M1  A1 [4]	any three correct; may be implied by 2 S's  correct sub in any correct S formula, ft $\Sigma s, \bar{x}, \bar{y}$  corr sub into 3 Ss and r, ft $\Sigma s, \bar{x}, \bar{y}$  cao	OR, using $S_{xx} = \Sigma(x - \bar{x})^2$ etc: $\bar{x} = \frac{1366}{5}$ or 273.2, $\bar{y} = \frac{17.6}{5}$ or 3.52, either: B1 $(-23.2)^2 + (-3.2)^2 + (-9.2)^2 + 16.8^2 + 18.8^2$ $0.68^2 + 0.18^2 + (-0.32)^2 + (-0.02)^2 + (-0.52)^2$ $(-23.2) \times 0.68 + (-3.2) \times 0.18 + (-9.2) \times (-0.32) + 16.8 \times (-0.02) + 18.8 \times (-0.52)$  If no working seen: -0.71: SC 3; -0.7: SC 1
(ii)	$b = \frac{-23.52}{1268.8}$ or $-\frac{147}{7930}$ or -0.0185 (3 sfs) $y - \frac{17.6}{5} = -0.0185(x - \frac{1366}{5})$ $\Rightarrow y = -0.019x + 8.6$ or better, ie 2 sfs enough $(y = -0.019 \times 280 + 8.6 \quad (= 3.39 \text{ to } 3.41))$  Est sales = £3390 to £3410 or 3.39 thousand to 3.41 thousand	M1 M1  A1  A1ft [4]	fit their $S_{xy}$ & $S_{xx}$ & $\Sigma s$ from (i) or $a = \frac{-17.6}{5} - (-0.0185) \times \frac{1366}{5}$ if a incorrect, must see method for M1 cao; must be "y = ..." coeffs that round to -0.019 & 8.6 to 2 sfs  fit their $y \times 1000$ , dep M1M1, dep sub 280 (not 280000) Allow "k" for thousand No working, ans in range: M1M1A0A1	use of x on y line: $b' = \frac{-23.52}{0.868}$ (or -27.1) M0 $x - \frac{1366}{5} = -27.1(y - \frac{17.6}{5})$ or $a' = \frac{1366}{5} - (-27.1) \times \frac{17.6}{5}$ M1 (if a' incorrect, must see method for M1) $x = -27.1y + 369$ cao A1  3277 or 3278 A0
(iii)	There may be other factors oe  Correlation does not imply causation oe	B1  [1]	or any suggestion of another factor that could be involved, eg Depends on state of the economy oe  Must state or clearly imply: EITHER <u>corr'n</u> does not imply <u>causation</u> OR there could be <u>another factor</u> involved  Ignore all else	NOT: Tourists & sales not nec'y linked Sales are not entirely dep on tourists Could be a coincidence Might be different other years More tourists wd incr sales -0.8 is not strong corr'n Only shows good neg corr'n Sample is small Could be affected by extremes Neg corr'n not nec'y imply neg relnship

3.

(i)	$H_0$ : A and B are not associated $H_1$ : A and B are associated	B1 [1]	For both. Allow indpt., not indpt.
(ii)	Yates $\chi^2 = \sum( O - E  - 0.5)^2/E$ which decreases the value	B1 B1 [2]	Dep '-0.5' seen.
(iii)	CV 5.024 seen 5.63 > CV and reject $H_0$ There is evidence at the 2 1/2% SL of an association between A and B	B1 M1 A1 [3]	Ft their CV Allow B1 if correct conclusion, but comparison not shown. CWO (ie from 5.024)

4.

first two $d'$ 's = $\pm 1$ $\Sigma d^2$ attempted (= 2) $1 - \frac{6 \times "2"}{7(7^2 - 1)}$ $= \frac{27}{28}$ or 0.964 (3 sfs)	B1 M1 M1 dep A1	$S_{xx}$ or $S_{yy} = 28$ B1 $S_{xy} = 27$ B1 $S_{xy} / \sqrt{(S_{xx}S_{yy})}$ M1 dep B1  1234567 & 1276543 (ans $^2/7$ ): MR, lose A1
--	--------------------------	--

5.

(i)	$\text{Mean} = (20+16+9)/75$ $= 0.6$ $3p = 0.6, p = 0.2$ AG	M1 A1 A1 3	
(ii)	$H_0: B(3, p)$ fits the data $(H_1: B(3, p)$ does not fit the data) Expected values 38.4 28.8 7.2 0.6  Combine last two cells $\chi^2 = 5.6^2/38.4 + 8.8^2/28.8 + 3.2^2/7.8$ $= 4.818$  $4.818 > 3.841$ Reject $H_0$ and conclude that there is insufficient evidence that $B(3, p)$ fits the data.	B1  M1 A1 A1 B1 M1 A1√ A1  B1√ M1 10	Or: $X \sim B(3, p)$ or $B(3, 0.2)$ Not 'Data fits model'  Use $B(3, 0.2) \times 75$ At least 2 correct All correct  With one correct At least 2 correct Ft E values Accept 4.82 cao  ft 4.818 SR1 If cells not combined: B1M1A1A1B0M1A1A0B1(5.991)M1 SR2: E-values rounded :B1M1A1A1 B1M1A1A0(4.865)B1M1
(iii)	$2.74 < 3.841$ , accept $H_0$ conclude that $B(6, p)$ fits the data	B1 1	Accept with no reason if evidence of method in (ii)

6.

7.

8.

9.

10.

11.

12.

13.

14.

15.

16.

17.

18.

19.

20.