

QQQ – Further Statistics 2 – Chapter 2 – Correlation (Version 1)

Total Marks: 23

(23 = Platinum, 21 = Gold, 18 = Silver, 16 = Bronze)

1. An estate agent in *Tornep* believes that houses further from the railway station are more expensive than those that are closer. She took a random sample of 22 three-bedroom houses in *Tornep* and calculated the product moment correlation coefficient between the house price and the distance from the station to be 0.3892

Stating your hypotheses clearly, use a 5% level of significance to test the estate agent's belief. State the critical region used in your test. **(3)**

(Total 3 marks)

Edexcel AS October 2020 Q1

2. Abena and Meghan are both given the same list of 10 films.

Each of them ranks the 10 films from most favourite to least favourite.

For the differences, d , between their ranks for these 10 films, $\sum d^2 = 84$

(a) Calculate Spearman's rank correlation coefficient between Abena's ranks and Meghan's ranks. **(1)**

A test is carried out at the 5% level of significance to see if there is agreement between their ranks for the films.

The hypotheses for the test are

$$H_0 : \rho_S = 0 \quad H_1 : \rho_S > 0$$

(b) (i) Find the critical region for the test.

(ii) State the conclusion of the test. **(2)**

An 11th film is added to the list. Abena and Meghan both agree that this film is their least favourite.

A new test is carried out at the 5% level of significance using the same hypotheses.

(c) Determine the conclusion of this test. You should state the test statistic and the critical value used. **(4)**

(Total 7 marks)

Edexcel AS June 2022 Q1

3. Phil measures the concentration of a radioactive element, c , and the amount of dissolved solids, a , of 8 random samples of groundwater. His results are shown in the table below.

Sample	A	B	C	D	E	F	G	H
c	625	700	650	645	720	600	825	665
a	1.28	1.30	1.00	1.20	1.55	1.15	1.40	1.45

Given that

$$S_{cc} = 34\,787.5 \quad S_{aa} = 0.217\,287\,5 \quad S_{ca} = 47.7625$$

- (a) calculate, to 3 decimal places, the product moment correlation coefficient between the concentration of the radioactive element and the amount of dissolved solids for these groundwater samples. (1)
- (b) Use your value of the product moment correlation coefficient to test whether or not there is evidence of a positive correlation between the concentration of this radioactive element and the amount of dissolved solids in groundwater. Use a 5% significance level. State your hypotheses clearly. (3)
- (c) Calculate, to 3 decimal places, Spearman's rank correlation coefficient between the concentration of the radioactive element and the amount of dissolved solids. (5)
- (d) Use your value of Spearman's rank correlation coefficient to test for evidence of a positive correlation between the concentration of the radioactive element and the amount of dissolved solids. Use a 5% significance level. State your hypotheses clearly. (3)
- (e) Using your conclusions in part (b) and part (d), comment on the possible relationship between these variables. (1)

(Total 13 marks)

Solutions

1.

Scheme	Marks	A0
$H_0 : \rho = 0 \quad H_1 : \rho > 0$	B1	2.5
5% one-tail cv for r is: 0.3598	M1	1.1b
Significant result so there is evidence to support the agent's belief	A1	2.2b
	(3 marks)	

Notes

B1 for both hypotheses correct in terms of ρ

M1 for use of tables to find the cv of 0.3598

A1 for a correct conclusion in context mentioning “belief” or description of this

2.

Mark scheme (a)

Scheme	Marks	A0
$\left[r_s = 1 - \frac{6 \times 84}{10(10^2 - 1)} \right]$ $= 0.4909 \dots \quad \text{awrt } \underline{\underline{0.491}}$	B1	1.1b
	(1)	

Notes

B1: awrt 0.491 (allow $\frac{27}{55}$)

Mark scheme (b)

Scheme	Marks	A0
(i) $(1 >) r_s > 0.5636$	B1	1.1b
(ii) (0.491 is not in the critical region) There is insufficient evidence of agreement between their film ranks.	Blft	2.2b

	(2)	
--	-----	--

Notes

(i) **B1**: Correct critical region with 0.5636 or better. Condone use of ρ instead of r_s

(ii) **B1ft**: Correct ft contextualised conclusion (must include film or ranks) based on their (a) and their CR Allow ft on their CV if a CR is not stated

Mark scheme (c)

Scheme	Marks	AO
new $r_s = 1 - \frac{6 \times 84}{11(11^2 - 1)}$	M1	1.1b
$= 0.61818 \dots$ awrt 0.618	A1	1.1b
new critical value is 0.5364	B1	1.1b
There is now sufficient evidence of agreement between their film ranks.	A1	2.2b
	(4)	
	(7 marks)	

Notes

M1: Use of formula with same $\sum d^2$ and 11

A1: awrt 0.618 (allow $\frac{34}{55}$)

B1: 0.5364 or better

A1: fully correct solution with awrt 0.618 and contextualised conclusion (must include film or ranks)

3.

Mark scheme (a)

Scheme	Marks
$r = \frac{S_{ca}}{\sqrt{S_{cc}S_{aa}}} = \frac{47.7625}{\sqrt{34787.5 \times 0.217287}} = 0.549361 \dots$	B1
	(1)

Notes

Mark scheme (b)

Scheme	Marks
$H_0 : \rho = 0, H_1 : \rho > 0$	B1
$(0.549 <) 0.6215$ (Not significant Insufficient evidence to reject H_0)	B1
Insufficient evidence of a positive correlation between the concentration of a radioactive element and the amount of dissolved solids in groundwater.	B1ft
	(3)

Notes

1st B1 Both correct. Require population parameter ρ and one tailed test.

2nd B1 cv 0.6215

3rd B1 Context required. Must mention **concentration of a** radioactive **element** and **amount of dissolved solids**

Mark scheme (c)

Scheme	Marks																														
<table><tr><th>Sample</th><th>A</th><th>B</th><th>C</th><th>D</th><th>E</th><th>F</th><th>G</th><th>H</th><th></th></tr><tr><td>c</td><td>2</td><td>6</td><td>4</td><td>3</td><td>7</td><td>1</td><td>8</td><td>5</td><td></td></tr><tr><td>a</td><td>4</td><td>5</td><td>1</td><td>3</td><td>8</td><td>2</td><td>6</td><td>7</td><td></td></tr></table>	Sample	A	B	C	D	E	F	G	H		c	2	6	4	3	7	1	8	5		a	4	5	1	3	8	2	6	7		M1
Sample	A	B	C	D	E	F	G	H																							
c	2	6	4	3	7	1	8	5																							
a	4	5	1	3	8	2	6	7																							
<table><tr><td>d</td><td>-2</td><td>1</td><td>3</td><td>0</td><td>-1</td><td>-1</td><td>2</td><td>-2</td><td></td></tr><tr><td>d²</td><td>4</td><td>1</td><td>9</td><td>0</td><td>1</td><td>1</td><td>4</td><td>4</td><td>24</td></tr></table>	d	-2	1	3	0	-1	-1	2	-2		d ²	4	1	9	0	1	1	4	4	24	M1A1										
d	-2	1	3	0	-1	-1	2	-2																							
d ²	4	1	9	0	1	1	4	4	24																						
Note Reverse ranks $\sum d^2 = 144$																															
$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} = 1 - \frac{6 \times 24}{8(64 - 1)} = 0.71428 \dots$	M1A1																														
	(5)																														

Notes

1st M1 for an attempt to rank the concentration of a radioactive element **and** the amount of dissolved solids with at least 4 correct for each variable. Allow reverse ranks.

2nd M1 for attempt at d^2 row

1st A1 all correct

3rd M1 for use of the correct formula and an attempt to rank, follow through their $\sum d^2$ if clearly stated
If answer is not correct, a correct expression is required.

A1 awrt 0.714

Mark scheme (d)

Scheme	Marks
$H_0 : \rho_s = 0, H_1 : \rho_s > 0$	B1
$(0.714 >) 0.6429$ (Significant. Reject H_0)	B1
Evidence of a positive correlation between the concentration of a radioactive element and the amount of dissolved solids in groundwater.	B1ft
	(3)

Notes

1st B1 for both hypotheses in terms of ρ , one tail H_1 . Allow use of ρ_s .

Alternative hypothesis compatible with their ranking.

2nd B1 for cv of 0.6429

3rd B1ft for a correct contextualised comment. Must mention **concentration of a radioactive element** and the **amount of dissolved solids**. Follow through their r_s and their cv (provided it is $|cv| < 1$)

Don't insist on the word "positive" for a one-tailed test.

Mark scheme (e)

Scheme	Marks
Results of tests suggest (monotonic) non-linear relationship or assumptions for PMCC breached i.e. not (joint) normal .	B1
	(1)
	(13 marks)

Notes

B1 for '**non-linear**' oe, or '**not normal**'