3.8 Fieldwork: geographic inquiry and statistics

Name: _____ Class: _____ Date: _____ **Total: 30 marks**

KEY WORDS hypothesis • null hypothesis • random sampling • systematic sampling • stratified sampling • Spearman's rank correlation • statistically significant • chi-squared test

Objective

Build the skills for the **fieldwork** internal assessment: a precise fieldwork question, a **hypothesis** and **null hypothesis**, sampling, methods, presenting data, **Spearman's rank correlation** and chi-squared tests, and an honest conclusion and evaluation.

You must be able to:

- write a precise fieldwork question with a hypothesis and a null hypothesis
- choose random, systematic or stratified sampling and justify the choice
- choose graphs and maps that suit the data
- calculate Spearman's rank correlation and test it for significance
- evaluate methods and suggest realistic improvements
- know the six assessment criteria and their marks

1 Key knowledge and worked examples

The inquiry

The fieldwork report is worth 25 % at SL and 20 % at HL: up to 2,500 words, marked out of 25, on **primary data** you collect yourself.

criterion	marks
A fieldwork question and geographic context	3
B method(s) of investigation	3
C quality and treatment of information collected	5
D written analysis	10
E conclusion	2
F formal requirements	2

exercises.lol exercises.com.cn

- **Hypothesis**: a testable statement ("pedestrian numbers fall as distance from the square increases"). **Null hypothesis**: there is no relationship.
- **Random sampling**: random numbers choose points; no bias from the student, but may miss small areas. **Systematic sampling**: every 100 m or every tenth person; even cover, but may match a regular pattern. **Stratified sampling**: samples in proportion to known zones; every zone represented, but you must know the zones.
- Present data with the right tool: scatter graph (two variables), bar chart or divided bar (categories), flow map (movements), choropleth map (values by area).

Spearman's rank

Rank both variables (1 = smallest in each), find d (difference in ranks) for each site, then

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

r_s near $+1$: strong positive; near -1: strong negative; near 0: no relationship. Compare the size of r_s with the **critical value** (5 % level, two-tailed): $n = 8$: 0.738; $n = 10$: 0.648. If it is bigger, reject the null hypothesis: the relationship is **statistically significant**. Correlation does not prove cause.

For counts in categories use a **chi-squared test** (for a 2×2 table, 1 degree of freedom, critical value 3.84 at 5 %).

Worked example. $n = 8$, $\sum d^2 = 158$: $r_s = 1 - \frac{6 \times 158}{8 \times 63} = 1 - \frac{948}{504} \approx -0.88$; $0.88 > 0.738$, so the negative relationship is significant.

2 Practice

2.1 Distinguish between a **hypothesis** and a **null hypothesis**. [2]

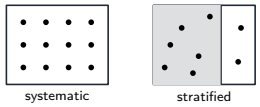
relationship — hypothesis

no relationship — null hypothesis

2.2 A student measured 10 sites and found $\sum d^2 = 30$. Calculate Spearman's rank correlation coefficient, and state whether the result is significant at the 5 % level (critical value 0.648). [3]

$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}$

2.3 Distinguish between **systematic** and **stratified** sampling. [2]



2.4 Suggest a suitable way to present each set of data: (A) pedestrian counts at 8 distances from a square; (B) the percentage of each land use along a street; (C) where 50 shoppers travelled from. [3]

A B C

2.5 Explain **two** ways a student could make pedestrian counts more reliable. [4]



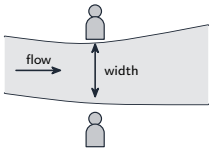
3 Exam-style questions

3.1 A student tested the hypothesis "river channel width increases with distance downstream". They measured width at 8 sites, every 500 m (systematic sampling).

site	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
distance downstream (m)	0	500	1,000	1,500	2,000	2,500	3,000	3,500
channel width (m)	1.2	1.5	1.4	2.1	2.6	2.4	3.3	3.8

(a) Write a null hypothesis for this investigation. [2]

(b) Calculate Spearman's rank correlation coefficient and state whether the result is significant at the 5 % level (critical value for $n = 8$: 0.738). [4]



3.2 Evaluate the methods the student used in 3.1, and suggest improvements. [10]

exercises.lol exercises.com.cn

Solutions

Marking: **M1** a correct method, **A1** the correct answer, **B1** an independent mark.

Each answer shows the steps, then the **result**.

2.1

hypothesis: a testable statement that a relationship exists, e.g. "river width increases downstream"; null hypothesis: the statement that there is **no** relationship, which a statistical test can reject

2.2

$r_s = 1 - \dfrac{6 \times 30}{10 \times 99} = 1 - \dfrac{180}{990} \approx \mathbf{0.82}$; $0.82 > 0.648$, so a significant positive relationship: reject the null hypothesis

2.3

systematic: samples at regular intervals, e.g. every 100 m along a street or every tenth person; stratified: the number of samples in each zone is in proportion to the size of the zone, so every zone is represented

2.4

A: scatter graph; B: divided bar chart or pie chart; C: desire-line or flow map

2.5

any two, explained: count at several times and on different days, so one unusual hour does not distort the result; repeat counts at each site and take the mean, which reduces random error; use the same length of count (e.g. 5 minutes) and a clicker counter at every site, so results are comparable; use more sites or several streets, so the pattern is not caused by one street's shops

3.1

(a) there is no relationship between channel width and distance downstream

(b) ranks of width (1 = narrowest): 1, 3, 2, 4, 6, 5, 7, 8; d: 0, −1, 1, 0, −1, 1, 0, 0; $\sum d^2 = 4$; $r_s = 1 - \dfrac{6 \times 4}{8 \times 63} = 1 - \dfrac{24}{504} \approx \mathbf{0.95}$; $0.95 > 0.738$: a significant positive relationship, so reject the null hypothesis

3.2

Mark with the markbands: [9-10] focused on the command term, detailed and accurate knowledge, well-developed explanation, specific named examples (places, dates, numbers), balanced evaluation and a supported conclusion; [7-8] mostly focused, good knowledge and examples, some evaluation; [5-6] relevant knowledge and examples, explanation partly developed, little evaluation; [3-4] general knowledge, few examples, mostly description; [1-2] little that is relevant. Points that may be included: strength: systematic sampling every 500 m covers the whole stretch evenly; the method is simple and easy to repeat; a statistical test gives an objective judgement; limitation: only one width measurement per site, so a narrow or wide spot can distort the result; only 8 sites; more sites would give a more reliable test; width depends on the water level, which changes with rain, so all sites should be measured on the same day in similar weather; bends, bridges or channel works can change width locally; improvements: measure three widths at each site and use the mean; add depth and velocity to calculate discharge; carry out a risk assessment and never enter deep or fast water; conclusion: a sound simple method whose reliability would improve with repeated measurements and more sites

exercises.lol exercises.com.cn