University of South Australia

2006 End of Year Examination

School	NBE	Subject Area & Catalogue number	GEOE 2020
Course Name		Spatial Statistics	

Official Reading Time: 10 Minutes
Writing Time: 2 Hours

Instructions to Candidates:

General Instructions to Candidates

Total marks = 100
Attempt question 1 and any other three questions.
All questions are of equal value.
Please ensure the front of answer books are completed with your name, student ID number, program, and course.

Additional Instructions

Read this examination question paper *very carefully*.
Clearly state any assumptions made.
Use examples and diagrams, where appropriate, in answering the questions.
Start a new question on a new page in the answer book.
Insert the number of the question at the top left hand margin on each page.
Insert the *main numbers* of the questions attempted on the front cover of the answer book.
Use both sides of the paper in the answer book.

NOTATION and ABBREVIATIONS

The terms "arithmetic mean" and "mean" are synonymous.
μ and σ Population mean and standard deviation.

Permitted Materials

- Any calculator

QUESTION 1

(a) What is it that distinguishes classical statistics from spatial statistics?

(b) On the basis of the *First Law of Geography*, viz., *"everything on the surface of the earth is related, but closer things are more related than things further apart"*, explain what is meant by the term **"spatial autocorrelation".**

(c) For which entities of vector spatial data, viz., *points, lines and polygons* may spatial autocorrelation be determined?

(d) What statistics may be used to measure spatial autocorrelation for the vector entities selected in part (c) of this question?

(e) Distinguish between local and global measures of spatial autocorrelation.

(f) With the aid of diagrams, demonstrate the meanings of strongly positive, near zero and strongly negative numerical values of spatial autocorrelation.

QUESTION 2

PART 1 (Part 2 appears at the top of page 3)

(a) What tests can be applied to a frequency distribution to determine whether or not it is normally distributed?

The *normal probability* distribution or *normal* distribution is defined by the *normal probability distribution function* or *normal probability function f(x)*, which, for the continuous random variable *X*, can take on specific values $X = x$ where

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \qquad \text{for} \quad -\infty < x < \infty$$

(b) With the aid of sketches, show how the shape of the normal distribution curve or *normal curve* varies with changes in mean and variance.

By applying the transformation given by the *standard normal variable, Z* where

$$Z = \frac{X-\mu}{\sigma} \qquad \text{which can take on any specific value} \qquad z_i = \frac{x_i - \mu}{\sigma},$$

any normal distribution may be converted into a standard normal distribution.

(c) What are the values of the mean, variance and standard deviation of the standard normal distribution?

(d) Hence, show that the standard normal distribution is given by:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \qquad \text{for} \quad -\infty < z < \infty$$

(e) What is the area under the normal curve whether it is standardised or not?

QUESTION 2 CONTINUED

PART 2

(a) Without performing any mathematical derivations but with the use of one or more appropriate diagrams, explain how a "line of best fit", linear or otherwise, may be obtained to approximate a set of (x,y) coordinate pairs representing a relationship, (not necessarily linear), between the x and y values.

(b) What are the extreme values of the sample correlation coefficient of a linear regression? Describe the nature of the relationship between the x and y variables for values of the sample correlation coefficient near the extreme values.

QUESTION 3

With the aid of examples and diagrams where appropriate, describe the following spatial statistics which apply to point patterns:

(a) Mean centre;

(b) Weighted mean centre;

(c) Median centre;

(d) Standard distance and

(e) Standard deviational ellipse.

QUESTION 4

(a) With the use of diagrams, give two examples of line patterns in vector spatial data.

(b) Explain what is meant by the terms "straight-line length" (Slength), "true length" (Tlength) and sinuosity in the context of a line pattern. What is the minimum value of sinuosity? Describe, with the aid of the diagram, the kind of chain, which gives rise to a value of sinuosity greater than the minimum.

(c) For two chains in a line pattern, the Slengths and the Tlengths are given in the following table. Compute the sinuosity of the two chains, and hence compare their shapes.

ID	Slength	Tlength
1	386.53	596.84
2	180.06	217.74

(d) Show, with the use of a simple example and diagram, why the arithmetic mean is an inappropriate statistic to describe the direction of a line pattern. What, then, may be used as the directional mean of a chain in a line pattern?

(e) What are the minimum and maximum values of circular variance of a line pattern? Describe the nature of line patterns with small and large circular variance.

QUESTION 5

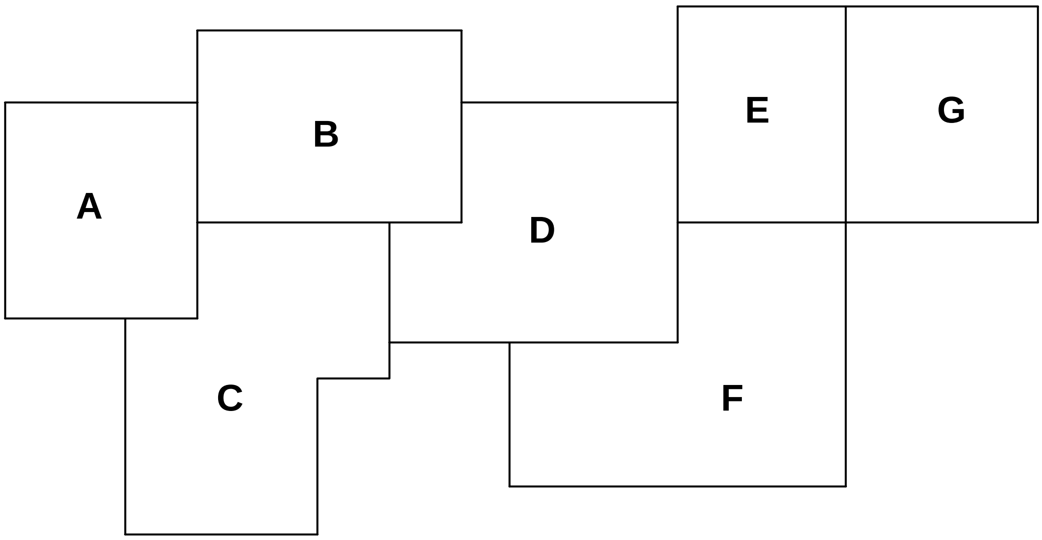
(a) What are spatial weights matrices, and what is their role in analysing polygon patterns?

(b) Briefly describe four different spatial weights matrices.

(c) The following diagram shows seven polygons named and labelled A to G. Construct the binary connectivity matrix of the seven polygons using the queen's case, showing the row sums and column sums. Perform a check to verify that the sum of each row agrees with the number of neighbours of the polygon represented by that row.

(d) What are the properties of this matrix?

(e) Hence, devise the stochastic or row-standardised matrix.



NOTE: There is no need to redraw this diagram in the answer book.