When interest extends beyond purely univariate considerations, there is no fixed rule governing the ordering of spatial samples for mapping and statistical analysis. Ratios and numerical differences have conventionally been the basis for comparing samples. As both these are known to be influenced by variations in sample size, they are unsuitable for the comparison of demographic samples which vary greatly in size. Where necessary, the effect of sample size can be accommodated by a modified form of the X^2 test statistic which is described in this paper.

The Signed Chi-square Measure for Mapping

M. Visvalingam

Census Research Unit, Department of Geography, University of Durham, Durham

INTRODUCTION

Ratios and numerical differences are among the most common forms of expression for the ordering and classification of data for bivariate mapping. They are considered to remove the effects of variations in the size of the base population and are therefore widely employed for inter-area, inter-group and temporal comparisons in demography and population geography.[1] Ratios, including such measures as proportions and percentages, are frequently encountered in both statistical and non-statistical work.

However, several writers have found the above standardisation procedures inadequate.[2–4] In general, numerical differences produce more extreme values in large populations while ratios have the opposite effect of tending to give the most extreme values in small populations. They exaggerate small differences in small populations; in addition, ratio measures may conceal large absolute differences in urban areas—differences which may be much more important because of the large numbers of people involved.

This paper suggests a different scheme for standardisation, namely the signed chi-square (X_S^2) measure which has been successfully used for the mapping of a wide range of themes.[8] It is a compromise measure which takes into account both the absolute and relative deviations from the norm, and is represented by a single cartographic symbol.

LIMITATIONS OF TRADITIONAL SCHEMES FOR ORDERING SAMPLES

Ratios and numerical differences have constituted the conventional means of ordering and scaling data for spatial units. Both these are influenced by variations in sample size. The sample size may be defined by the number of residents, of households or even of some other base population such as the number of males of a particular age group within the spatial unit. It conventionally forms the denominator of the ratio measure. That part of the sample which is of substantive interest, for example the number of children or of unemployed males, often (but not always) forms the numerator of the ratio measure. The difficulties associated with the use of ratios and numerical differences stems from two features: firstly, the distribution of sample sizes is often not constant. The variation in sample sizes is a marked feature in high resolution grid-square data, but it is also present in data for irregular areas such as Local Authority Areas. This is illustrated by the Lorenz curves in *Figure 1*, in which the cumulative percentage of total population is plotted against the cumulative percentage of data units, which had been ranked by population size. Moreover, often the section of the sample of substantive interest is constrained by the sample size and is closely related to it. When there is a high correlation between the size of the sample (denominator) and the group of interest (the numerator), the spatial distribution of the latter closely reflects the distribution of the former and hence is of little interest.

A second complication is that the bivariate relationship between denominator and numerator—or for that matter between the residual population (if applicable) and the numerator—lacks homoscedasticity. Numerical differences, being constrained by sample size, tend to be more extreme in large samples than in small ones. On the other hand, in small samples, very small differences can produce greatly different and often extreme ratios. In the extreme example of a kilometre grid square with one person,[5] only values of 0 and 100 *per cent* are possible; and in general, important variations in large populations tend to produce near-average ratios. The inherent assumption in the use of
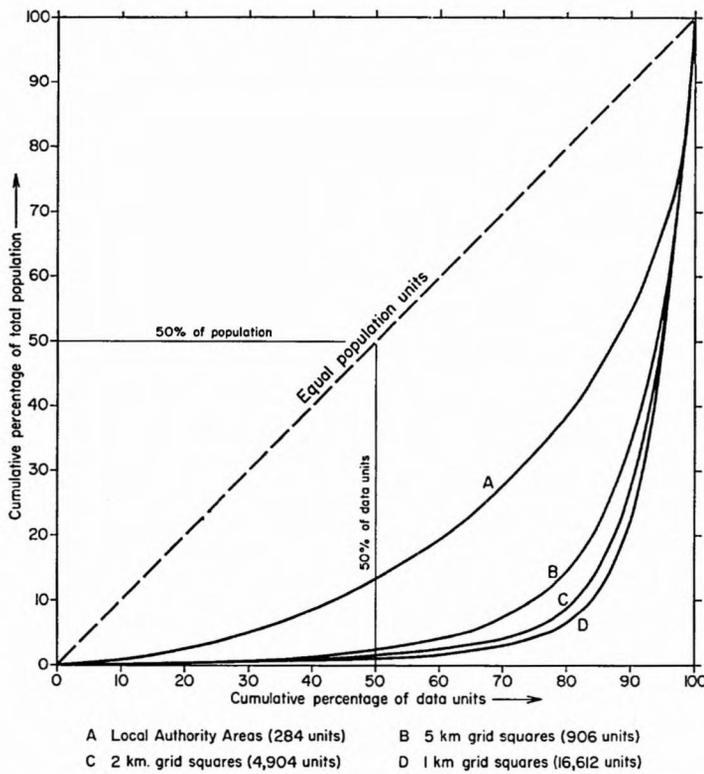
Figure 1. Lorenz curves showing the distribution of population among data units at four levels of aggregation.

ratio cut-offs is that the magnitude of deviation from expectation is a linear function of sample size as shown in Table 1. Moreover, the absence of a particular group is often more unusual in a large sample than in a small one, but the spatial units concerned are all given the same status by a ratio value of 0 *per cent*.

TABLE 1

Some implications of using a cut-off value of 60 per cent males

Total population (N)	Males	Deviation from an expectation of 48.5% males	Numerical differences between sexes
10	6	1.15	2
100	60	11.50	20
1,000	600	115.00	200
10,000	6,000	1,150.00	2,000
100,000	60,000	11,500.00	20,000

ALTERNATIVE SCHEMES

Several alternative or corrective measures have been suggested and tried. Dewdney and Rhind[3] considered the suppression of data for populations below a threshold value. In some of their maps, they removed the worst anomalies by considering only those one-kilometre squares with more than ten people—but the discussion so far makes it clear that a distorted ranking still exists. An alternative—the separate mapping and publication of maps of both absolute number and ratio measures of selected variables is not a satisfactory solution. Apart from doubling the number of maps, it produces a complex visual cor-

relation problem for the map reader especially if highly misleading ratio values still exist. The tagging on of significance levels to ratio values[2,6,10] is again not a viable solution for the mapping of high resolution data and does not substantially aid the recognition of spatial patterns on ratio maps.

The above measures attempt to redress a distorted ordination by what amounts to cosmetic procedures. The signed chi-square measure constitutes a more direct solution since it results in an ordering scheme which is intermediate to ratios, on the one hand, and absolute deviations on the other.

THE SIGNED CHI-SQUARE MEASURE

The signed chi-square measure is a modified form of the standard X^2 test statistic. X^2 measures the magnitude of the (absolute X relative) departure of observed values from an expected value.

$$X^2 = \Sigma \left[(O-E) \times \left(\frac{O-E}{E} \right) \right] = \Sigma \left(\frac{(O-E)^2}{E} \right) \quad (1)$$

where O is the observed frequency and E is the expected frequency.

Thus it is a compromise measure which simultaneously considers the magnitude of absolute and relative deviations from expectation. When there are two categories A and B in a sample of N, the above expression takes the following form:

$$X^2 = \frac{(O_A - E_A)^2}{E_A} + \frac{(O_B - E_B)^2}{E_B} \quad (2)$$

The X^2 measure is uni-directional. To distinguish between areas with an excess of A as opposed to B, the X^2 formula was modified as follows:

$$X_S^2 = \text{sgn} \, (O_A - E_A) \, X^2 \quad (3)$$

where sgn (x) is the signum function whose value is 1 for $x > 0$, -1 for $x < 0$ and O for $x = 0$. Thus areas with an excess and deficit of A are distinguished by positive and negative values of X_S^2 respectively. The sign is only a code and the numerical values remain exactly the same as the standard X^2 sample value.

If we nominally assume that expected values for A = B, then some further implications of X_S^2 standardisation become apparent. In this particular case:

$$X^2 = 2 \frac{(O_A - E_A)^2}{E_A} \quad (4)$$

From the above equation it can be appreciated that the same numerical deviation is rated more important in a smaller population than in a large one since the denominator E_A is smaller. If the above expression is inverted, we can derive O_A as:

$$O_A = E_A \pm \sqrt{\frac{X^2}{2}} \quad E_A = E_A \pm \text{tolerance} \quad (5)$$

Hence for a specified value of X^2, the tolerance is basically a square root function of the expected value and consequently of sample size. This corrects one of the difficulties associated with the use of ratio measures, namely that ratio measures assume a linear relationship between the magnitude of deviation from expectation and sample size.

A further understanding of the effect of the X_S^2 measure can be gained through its equivalent in the two-category case namely the Z-score measure.

$$Z = \frac{P-p}{\sqrt{pq/N}} \qquad (6A)$$

and

$$Z^2 = \frac{(P-p)^2}{pq/N} = X^2 \qquad (6B)$$

where p is the population proportion in Category A, *i.e.* the expected proportion,

q is the population proportion in Category B.

P is the proportion of Category A in the sample, *i.e.* the observed proportion,

N is the sample size,

$p+q = 1.0$, and

$\sqrt{pq/N}$ is the standard error for proportions.

Thus it can be appreciated that in both the X^2 and Z-score ordering schemes, the departure of the observed ratios are expressed in units of standard errors, *i.e.* in terms of the inherent variability which is inversely proportional to the sample size. Thus ratio values are standardised with respect to sample size and unless observed ratios coincide with expected values, the same sample value is deemed more extreme in a larger population than in a smaller one.

The Z-score and X_S^2 methods of computation are equally valid for proportions or the two-category case. The X_S^2 model is favoured because it is more general. The number of terms in the expression may be expanded or contracted for other applications. Only one term was used for standardising open ratios, *i.e.* when the numerator and denominator were two separate samples. The value of such ratios can be in excess of unity, for example in the ratio of children born to young women. The X_S^2 map of the fertility of young women shows marked regional contrasts.[8] On the other hand, several terms could be used for the study of age composition[8] or the mix of housing types or even the mix of socio-economic groups.

X_S^2 values may be mapped by proportional symbols. However, for maps of a small scale, it may be necessary to classify the data and adopt some sort of shading scheme. The extreme 10 *per cent* of cases are mapped in the examples given in this paper (see below). While a quantile system for the delimitation of class boundaries was excellent for the comparison of the ratio and X_S^2 ordering schemes, cut-off values are preferable in general as shown by Visvalingam and Dewdney.[5] Conventionally, X^2 values are associated with some level of statistical significance and X_S^2 cut-off values of ± 3.84 have been used successfully.

Some caution must be exercised in using the probability levels too rigidly. It has been suggested (J. Besag, personal communication) that demographic samples may not possess the necessary condition of statistical independence. Further as stated by White, 'significance may be measured only in terms of some *a priori* expectation'.[9] The key issue in the use of X_S^2 is the formulation of expectation. Values for X^2 calculations are usually derived mechanically from contingency tables or related to some probability distribution such as the Binomial or Poisson[2] or Normal[10] distributions. The population proportion, p, often substitutes for

expectation in the calculation of Z-score. Experience with several variables in the Census Research Unit of Durham suggests that the use of national averages may be adequate for most general purpose maps.

COMPARISON OF RATIO AND SIGNED CHI-SQUARE ORDERING SCHEMES

The illustrations in this paper refer to the entire land area within the three 100 kilometre squares whose south west corners are defined by grid references 300 400, 200 300 and 300 300 respectively; this covers most of Lancashire, Cheshire and North Wales. This section of the country was chosen for several reasons. It is sufficiently large to display a variety of general conditions. The settlements contained have a wide range of population sizes and are known to possess a marked spatial sorting of males and females, this was an important consideration since the two schemes for ordering data are compared using these variables. The data for the Local Authority Areas (LAA) were extracted from the appropriate published census volumes by J. C. Dewdney, while the data at the one-kilometre grid square level were provided on magnetic tapes by the Office of Population Censuses and Surveys.

Maps provide a convenient means for illustrating the relative performance of ratio and X_S^2 ordering schemes. Only the two extreme 10 *per cent* of cases are shown. The quantile system, was considered to offer the most objective basis for the delimitation of extreme cases for purposes of identifying the disparities between the two ordering schemes. The decile system was an arbitrary choice.

Figures 2 and *3* are based on data for 2×2 kilometre grid squares; there are 4940 grid squares of this resolution in the study area. The populations in these areas range from a single person to 46,914 persons per data unit and the frequency distribution of population sizes can be seen in *Figure 1*.

Figure 2 shows the extremes of masculinity as defined by the masculinity proportion, *i.e.* the proportion of males in the total population. *Figure 3* shows the spatial distribution of masculinity as defined by X_S^2 (refer equations 2 and 3). The X_S^2 values were calculated using an expectation based on the 1971 national average, *i.e.* 48.5 *per cent* males.

It is near impossible to identify any marked spatial trends in *Figure 2*, in which the punctiform pattern of extreme ratio values are generally associated with small populations in inland rural Wales and Lancashire; there are extensive blanks in the south-east of the study area and in the Liverpool and Manchester conurbations. The spatial pattern of extreme X_S^2 values in *Figure 3* exhibits contiguity and is more readily interpreted. Major blocks of contiguous squares with low masculinity are predominantly coastal and pick out holiday resorts and retirement areas. Other clusters of low masculinity squares occur in the inner suburbs of Liverpool, Manchester and several smaller towns. High masculinity occurs at both ends of the rural–urban continuum, being particularly marked in the Potteries, Wolverhampton and Walsall and in the belt extending from central Manchester, towards Liverpool. It is not only associated with urban–industrial locations and recent large-scale housing development but also with military establishments and prisons (a more extensive discussion is provided by Visvalingam and Dewdney[5]).

The disparity between the ratio and X_S^2 ordering schemes is even more noticeable at the one kilometre level, as shown
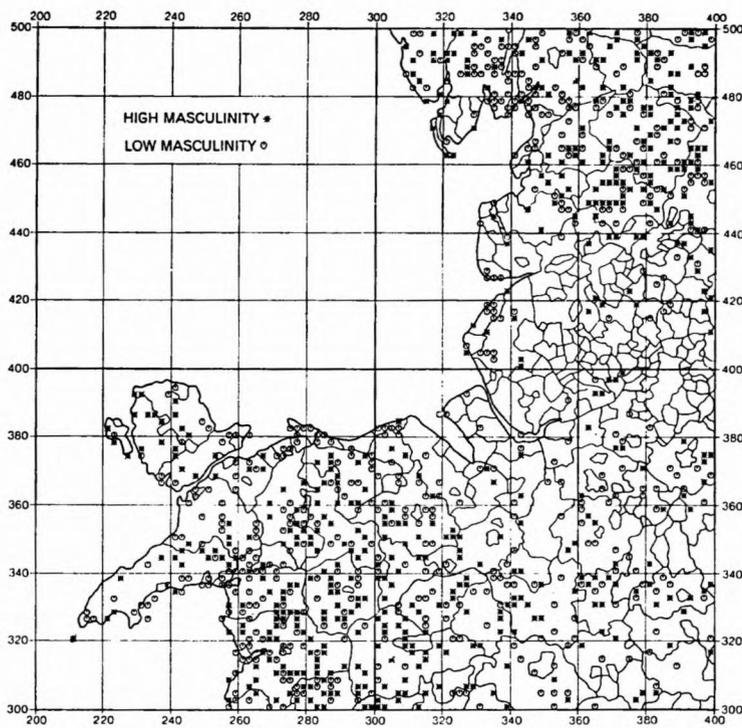
Figure 2. Extremes of masculinity as indicated by the masculinity proportion in 2 × 2 km grid squares.
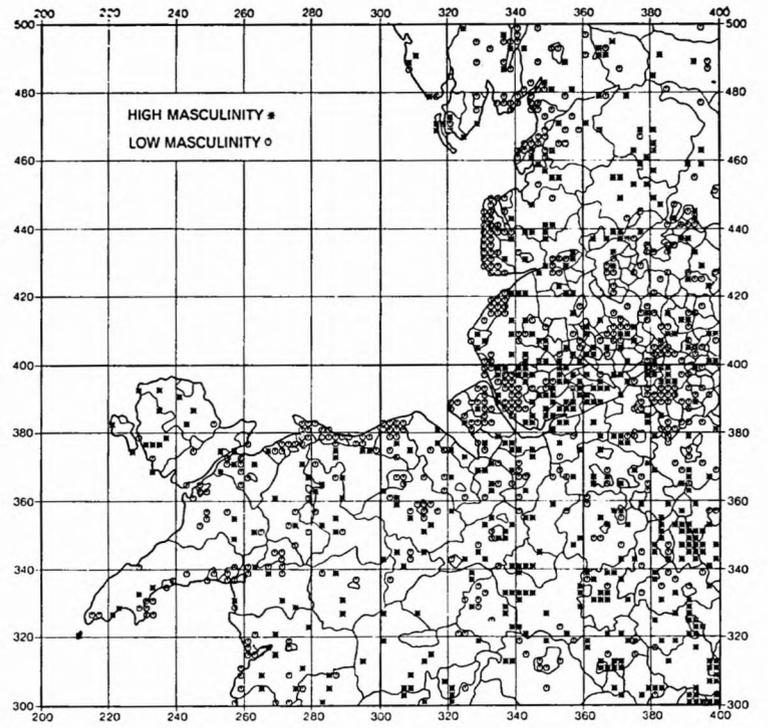
Figure 3. Extremes of masculinity as indicated by the X_S^2 measure in 2 × 2 km grid squares.
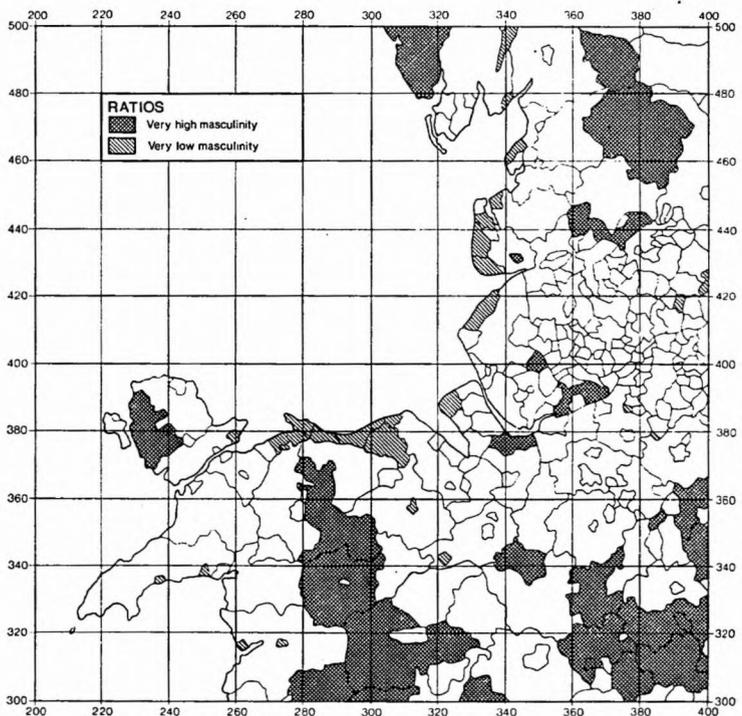
Ratio map showing the ten per cent of LAAs with highest and lowest masculinity

LAAs with masculinity ≥ 49.99%
No. of units = 28; proportion of total units = 10.0%

LAAs with masculinity ≤ 46.48%
No. of units = 28; proportion of total units = 10.0%

Figure 4. Ratio map showing the 10 *per cent* of LAAs with highest and lowest values.
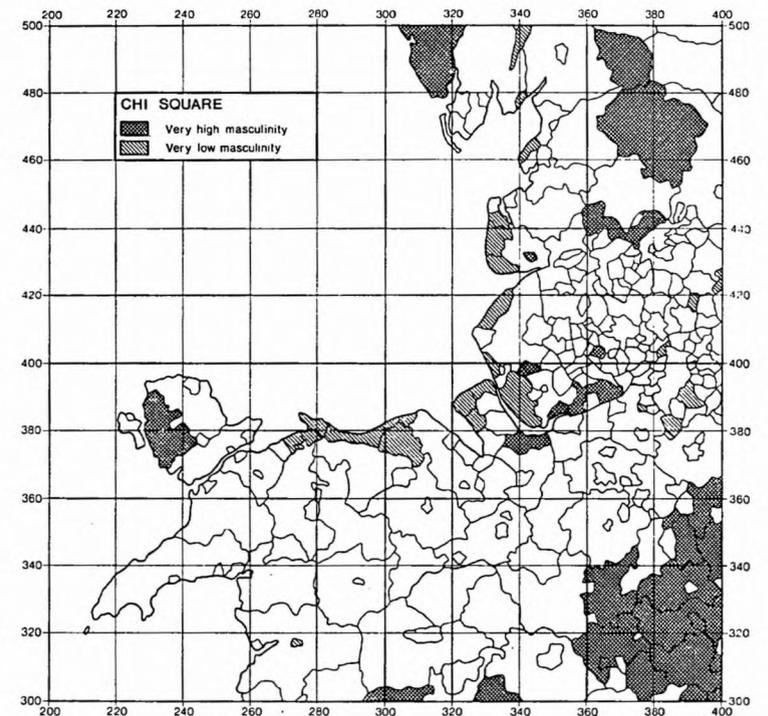
X_S^2 map showing the ten per cent of LAAs with highest and lowest masculinity

LAAs with X_S^2 ≥ 18.820
No. of units = 28; proportion of total units = 10.0%

LAAs with X_S^2 ≤ - 20.523
No. of units = 28; proportion of total units = 10.0%

Figure 5. X_S^2 map showing the 10 *per cent* of LAAs with highest and lowest values.

by the above authors. The X_S^2 ordering scheme yielded consistent spatial patterns of extremes of masculinity at the 1 × 1, 2 × 2 and 5 × 5 kilometre square levels. The ratio system on the other hand is not as robust. While the spatial patterns of masculinity as portrayed by ratio and X_S^2 measures are markedly different at the 1 × 1 and 2 × 2 kilometre square levels, the ratio maps begin to reflect the spatial pattern portrayed by X_S^2 values at coarser levels of aggregation, especially at the LAA level as shown in *Figures 4* and *5*.

It must be stressed that the difficulty associated with the use of ratio values does not vanish with an increase in the size of areal units or with the size of the sample population. For the problem is not a function of sample size (although small samples are always unreliable) so much as the incomparability of sample sizes. It could be solved if the spatial units are carved up to yield equal population samples. While irregular administrative areas are intended to produce more comparable samples, they are by no means equal or even near-equal population units; the

TABLE 2

The 10 Highest and Lowest Ratio and X_S^2 Values for Local Authority Areas

	Area and location		Population	Males	Ratio (% males)	X_S^2 value
	Ratio 10 highest values					
1	Shifnal R. D.	(377 307)	15,869	8,587	54.112	200.082
2	Sedbergh R. D.	(371 491)	3,544	1,912	53.950	42.150
3	Clitheroe R. D.	(368 440)	9,456	5,075	53.670	101.177
4	Kirkham U. D.	(343 432)	6,436	3,405	52.906	50.011
5	Penllyn R. D.	(289 334)	2,313	1,205	52.097	11.980
6	Clun & Bishop's Castle R. D.	(333 301)	8,869	4,577	51.607	34.272
7	Machynlleth R. D.	(285 305)	2,502	1,291	51.599	9.618
8	Maelor R. D.	(345 341)	4,679	2,393	51.143	13.090
9	Newton and Llanidloes R. D.	(306 301)	8,223	4,203	51.113	22.473
10	Millon R. D.	(315 492)	14,089	7,195	51.068	37.204
	Ratio 10 lowest values					
1	Grange U. D.	(341 478)	3,474	1,395	40.155	−96.847
2	Dolgellau U. D.	(274 317)	2,567	1,062	41.371	−52.228
3	Criccieth U. D.	(251 339)	1,505	646	42.924	−18.737
4	Llandudno U. D.	(280 381)	19,077	8,295	43.482	−192.343
5	Colwyn Bay M. B.	(287 378)	25,564	11,162	43.663	−239.461
6	Lytham St. Annes M. B.	(334 429)	40,299	17,765	44.083	−314.776
7	Windermere U. D.	(341 497)	8,065	3,557	44.104	−62.394
8	Barmouth U. D.	(263 317)	2,106	929	44.112	−16.234
9	Prestatyn U. D.	(306 383)	14,515	6,462	44.519	−92.077
10	Penmaenmawr U. D.	(273 376)	3,991	1,787	44.776	−22.162
	X_S^2 10 highest values					
1	Shifnal R. D.	(377 307)	15,869	8,587	54.112	200.082
2	Wolverhampton C. B.	(392 301)	269,112	133,976	49.784	177.768
3	Bolton C. B.	(370 410)	154,199	74,709	50.073	147.876
4	Warrington R. D.	(364 391)	50,420	25,684	50.940	120.191
5	Clitheroe R. D.	(368 440)	9,456	5,075	53.670	101.177
6	Walsall C. B.	(399 301)	184,738	91,605	49.588	87.476
7	Stafford M. B.	(393 322)	55,001	27,652	50.275	69.413
8	Cannock U. D.	(399 310)	55,882	28,022	50.145	60.538
9	Wellington R. D.	(367 317)	30,297	15,360	50.698	58.606
10	Ellesmere Pt. M. B.	(340 376)	61,637	30,831	50.020	57.035
	X_S^2 10 lowest values					
1	Blackpool C. B.	(330 436)	151,860	69,338	45.659	−490.659
2	Southport C. B.	(333 416)	84,574	37,921	44.838	−454.155
3	Lytham St. Annes M. B.	(334 429)	40,299	17,765	44.083	−314.776
4	Colwyn Bay M. B.	(287 378)	25,564	11,162	43.663	−239.461
5	Morecambe M. B.	(343 463)	41,908	18,848	44.975	−208.514
6	Llandudno U. D.	(280 381)	19,077	8,295	43.482	−192.343
7	Grange U. D.	(341 478)	3,474	1,395	40.155	−96.847
8	Prestatyn U. D.	(306 383)	14,515	6,462	44.519	−92.077
9	Hoylake U. D.	(324 388)	32,277	14,802	45.859	−90.112
10	Crosby M. B.	(332 400)	57,497	26,756	46.535	−88.919

frequency distribution of population sizes in the LAAs is still not uniform as shown in *Figure 1*. The LAAs included in *Figures 4* and *5* have population sizes which range from 729 to 610,113 persons. Since the population base is quite large, masculinity proportions are not as wildly variable as in the 1×1 and 2×2 grid square levels, but are confined to values within the range 40.15 to 54.11 *per cent* males. Despite this, the problems associated with ratio measures are present even at the LAA level. This is manifested to some extent in *Figures 4* and *5* but become more apparent when the LAAs are ranked by their ratio and X_S^2 values, refer Table 2. The relatively smaller populations continue to dominate the extremes. Hence, the present discussion is relevant to users of all types of spatial data, except when the spatial units were derived to yield equal or near-equal population samples.

There will be situations when the mapping of absolute numbers, absolute deviations and ratio expressions are particularly relevant. However, experience at the Census Research Unit has shown that the X_S^2 measure can be successfully applied over a wide range of variables and has been found particularly useful for highlighting contiguity relationships.

CONCLUSION

In conclusion, the limitations of using X_S^2 maps must be considered along with the ramifications for other studies. While X_S^2 maps are excellent for picking out spatial patterns, they cannot be used in the conventional manner in which ratio and absolute number maps are used for deriving quantitative information. A X_S^2 value summarises a relationship between ratios and absolute numbers; and while values of X_S^2 may be read-off the map, the latter do not contain any clue to the particular ratio or absolute numbers involved. Moreover, no external factors are involved in the derivation of absolute number or ratio maps. Thus, while the appearance of these maps may change with alterations to the number of classes or class boundaries, the order of data units remains static and unchanged and implies a single solution to multiple problems. The value for expectation used in the calculation of X_S^2 values has some effect on the relative order of data units. Thus the proper use of X_S^2 demands that the projection of expectation corresponds to the object of the academic exercise as ranking is relative to some *a priori* expectation.

The use of ratios is not confined to mapping. Multivariate exercises commonly include ratio variables and there has been much discussion concerning the choice of suitable transformations and scaling procedures.[11] The problems of ordering in these exercises are even more crucial and fundamental than those of scaling. Ratios based on small populations or highly varying base populations are unsuitable for multivariate analysis since the extremes, which are the products of small populations, are bound to dominate the calculation of correlation coefficients and set the trend in regressions. Although maps have been used to illustrate the disparities between ratio and X_S^2 ordering schemes, it must be emphasised that the implications of this work extend beyond cartographic concerns.

REFERENCES

1. H. S. Shryock, J. S. Siegel, and Associates, *Methods and Materials of Demography*, U.S. Govt. Printing Office, Washington, D.C., Vol. 1, 1973, 188–200.
2. M. Choynowski, Maps based on probabilities, *J. Am. Statist. Ass.* **54** (1959), 385–388.
3. J. C. Dewdney and D. W. Rhind, (eds.), *People in Durham—a census atlas*, Census Research Unit, Dept. of Geography, University of Durham, 1975, pp. 34–35.
4. M. Visvalingam, Chi-square as an alternative to ratios for statistical mapping and analysis, Census Research Unit, *Working Paper No. 8*, Dept. of Geography, Durham, 1976, p. 21.
5. M. Visvalingam, and J. C. Dewdney, The effects of the size of areal units on ratio and chi-square mapping, Census Research Unit, *Working Paper No. 10*, Dept. of Geography, Durham, 1977, p. 31.
6. G. M. Howe, *National Atlas of Disease mortality in the United Kingdom*, Nelson, London, 1970.
7. M. R. Spiegel, *Theory and problems of statistics*, Schaum's outline series, McGraw-Hill, New York, 1961, pp. 170 and 208.
8. Census Research Unit, *People in Britain—a census atlas*, H.M.S.O., forthcoming.
9. R. R. White, Probability maps of leukaemia mortalities in England and Wales, *Medical Geography Techniques and Field Studies* (ed. N. D. McGlashan), Methuen, London, 1972, pp. 173–85.
10. G. B. Norcliffe, Probability mapping of growth processes, *Economic Geography*, **48** (1972), 428–38.
11. I. S. Evans, J. W. Catterall, and D. W. Rhind, Specific transformations are necessary, Census Research Unit, *Working Paper No. 2*, Dept. of Geography, Durham, 1975, p. 12.