Methods of Studying Population Distribution

1.1 Introduction

Till recently, demographers did not give much importance to the study of distribution of populations. With realization of its crucial significance to understanding of population trends, various methods for studying population distribution—cartographic as well as statistical—have been developed. The basic information used in most studies is the census enumeration by geographic sub-divisions of a country, or other territorial units.

The spatial distribution of populations and settlements across a country and their interconnectivity and accessibility from urban centres are important for delivering healthcare, distributing resources and economic development. Three spatial features influence the economic development of a region: the density (e.g. agglomeration, scale economies), the distance (e.g. spatial mobility and access) and division (e.g. the spatial integration of economies). Improving access to people and markets is a key driver for development and plays an important role in poverty reduction. Development among populations depends on access to markets for buying and selling goods, to water and fuel, and to various social and economic services such as education, healthcare or banking and credit. The lack of reliable transport system forces rural populations to spend a significant amount of time in travelling to meet basic needs and increases the transport costs incurred to access these services. These factors often mean that isolation is seen as the main contributor to poverty according to the poor people themselves. The proximity of a major settlement provides business for isolated populations, and connectivity with international and regional markets creates economic opportunities. Accessing populations efficiently is also of key importance in public health, for delivering equitable and complete healthcare, for planning vaccine campaigns or distributing resources. The measurement of accessibility of populations and settlements is therefore of importance in measuring progress towards achieving these goals.

1.2 Cartographic Methods

Cartographic methods are, essentially, representation of population size on maps. Maps are important visual tools for the studying population size. Through maps, patterns of population distribution can be described easily and effectively. There are two main kinds of thematic maps, symbol maps and line maps. These days, GIS (Geographical Information System) techniques are used to prepare several types of maps:

1.2.1 Symbol maps: Symbol maps are used to portray the distribution of a characteristic either at different points, or within specified areas. The following symbolic maps are used for depicting population distribution.

1.2.2 Choropleth maps: The distribution of population in different areas can be shown by different shades on a map. This is called a choropleth map. Data are grouped into classes and each class is assigned a shade. The intensity of shade reflects the intensity of the category it is meant to portray. Since these maps use the administrative boundaries (e.g., districts), population variations within the administrative unit are not reflected by them.

1.2.3 Dot maps: The simplest symbol which appears to locate to a point is the dot. Each dot represents a certain number of persons. To express the population of an area (e.g., a village), the dots are placed regularly throughout the area, or as close to the actual settlements as possible. Obviously, the latter is the better alternative. The problem arises when there is an excessive clustering of dots in areas where large numbers need to be portrayed and, at the other extreme, where very small numbers are to be portrayed, they may not be sufficient to be represented by a dot. Therefore, the scale for using dots should be selected very carefully.

1.2.4 Areal and volumetric symbols: Symbols like circles, spheres, squares, cubes etc. are widely used to show the spatial variation of a population, particularly when the populations of different cities are to be shown. For a large difference between the smallest and the biggest units, one should use volumetric symbols (population size represented by volume) like spheres and cubes; but when the variation is not much, areal symbols (population size represented by area) like circles and squares should be used.

The symbols are placed at the appropriate location of the city on the map. Their scale must be chosen very carefully with respect to the scale of map; otherwise interpretation of the maps becomes difficult.

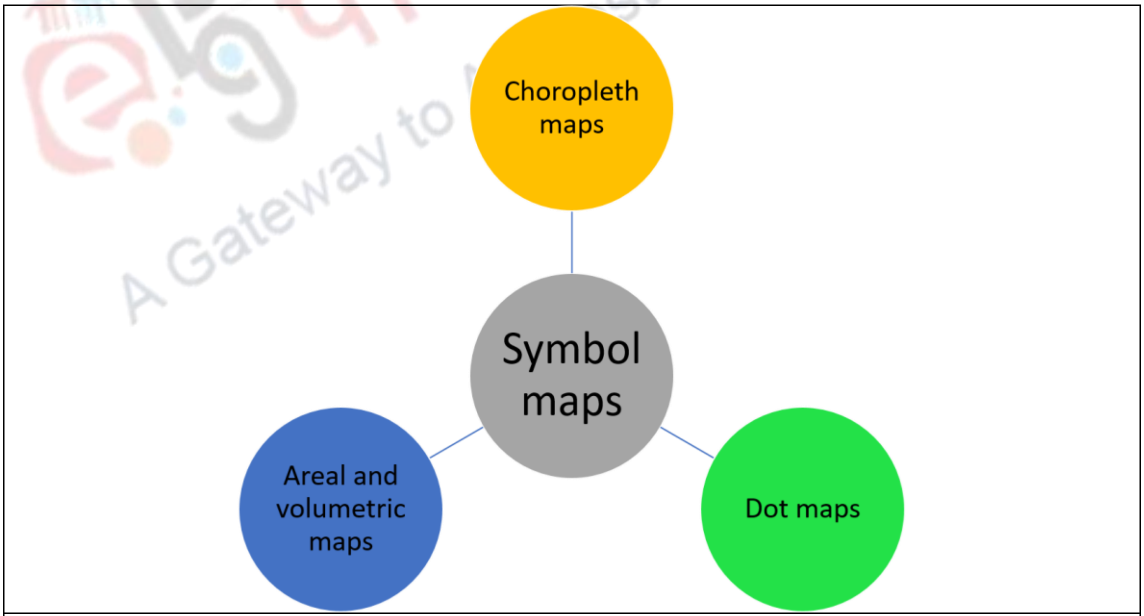

Figure 1: Types of Symbol Maps

Source: Data above

1.2.5 Line Maps: Line maps are used to show various relations between points or areas. The flow of migrants, traffic, commodities, etc. can be shown by line maps. Two types of line maps are briefly described here:

1.2.6 Routed Flows: When the actual route between the two points is known, flow lines can be constructed along that route. Their width is proportional to the volume of the flow at every point on the route. Scales are used by taking the simple values and not the square roots or log values which distort the visual impression, making interpretation difficult.

1.2.7 Non-routed Flows: These maps show the linkages between the different points in the diagrammatic form as straight or curved lines, either drawn in thickness proportional to the volume of the flow, or simply as symbolic links with no reference to the amount of flow. Usually the patterns of migratory movements are shown through such maps. Where the volume of flow, or the number of routes, is large, interpretation of such maps becomes rather difficult because of the criss-crossing and overlapping of so many lines.

The basic problem with cartographic methods is that population distribution data are usually available only by administrative and political divisions. As distributional representation is homogeneous, it does capture areas of concentration or dispersion. Also, areas with no human habitation, such as mountains, rivers or deserts, are also represented with population.

1.3 Graphic techniques: When the distribution of population is to be shown for the same place over a period, graphic techniques for frequency distribution can also be used to represent population distribution. The most commonly used techniques are bar diagrams and line graphs.

1.3.1 Bar diagrams: Bar diagrams are usually used to show the class frequency plotted against the class. The year of reference is mentioned on the X-axis (or abscissa), and the population for a given year on the y-axis (or ordinate). The scale can be decided by considering the minimum and maximum values. This technique is used for non-continuous (discrete) data.

1.3.2 Line graph: Sometimes, the data are plotted through continuous lines, which are known as *frequency curves*. These lines can clearly show the variations in population concentration or any characteristic of a population in a place over time. They are particularly useful when the temporal variation is to be presented for several places simultaneously.

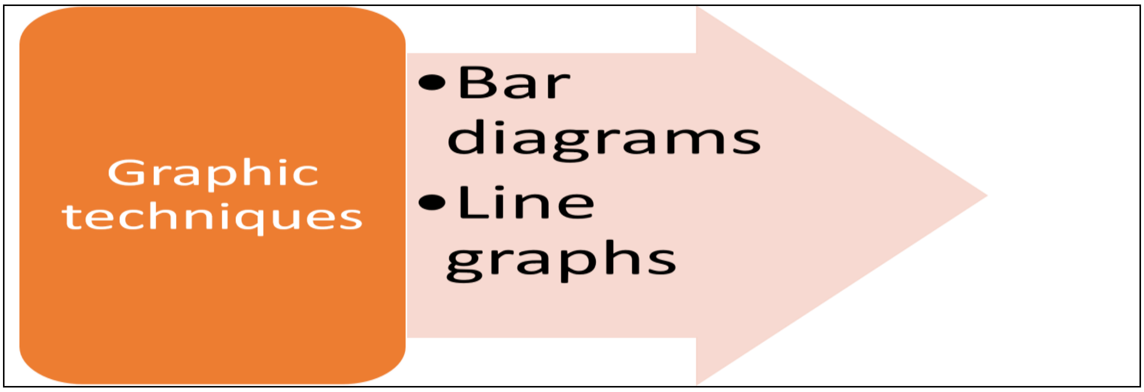

Figure 2: Types of Graphic techniques commonly used to represent population distribution
Source: Data above

1.4 Statistical Methods: Certain statistical techniques are used for measuring the spatial distribution of population and comparing the distribution in one area with another. We will discuss some of the most common techniques used in this regard.

1.4.1 Density: As mentioned earlier, the most common expression for measuring spatial distribution of population is the density. Various types of density are used. The *Arithmetic Density*, or *Crude Density*, is calculated by dividing the total number of people in a region with its area. Unless otherwise specified, density always refers to the crude arithmetic density. However, it does not fully explain the variations in human occupancy within a region. For example, the population density of Egypt was 43 persons per square km. in 1982, but nearly 95 per cent of the population was concentrated in the Nile river valley and delta.

To relate population to agricultural land, another expression is used, which is called *Physiological Density*. It is the number of people per unit area of land suitable for agriculture. Physiological Density is a more sensitive measure and reflects the demand of the people for land to grow their food.

Another type of density is the *Agricultural Density*, which is the ratio of the number of farmers to the total farm land. This reflects the efficiency of a society's agriculture practices. Agricultural density is low in economically advanced societies, where a small number of farmers can cultivate a large area of land. This is because of the high level of mechanization in agriculture.

1.4.2 Lorenz Curve: A Lorenz Curve is used to show the population-space relationship of concentration and dispersion. The actual distribution of population in space is plotted against the ideal or theoretical distribution, and the unevenness of population is traced. The curve was developed in 1905 by Lorenz, after whom it is today known as, to measure the inequalities in the distribution of wealth.

To plot the Lorenz curve, the size of population in each urban locality is required. It is arranged in an order such that the locality of highest population is the leading one. If there are too many urban areas, these can be grouped into different classes and then arranged in an ascending/descending order.

The proportional distribution of urban population and the number of localities are calculated. Cumulative proportion of urban population x_i and cumulative proportion of urban localities y_i are calculated. The value y_i is plotted against x_i and a smooth curve is drawn. For comparison, a diagonal line is drawn at 45^0 to show the condition of equal distribution. The curve should follow the diagonal if urban population and number of localities are evenly distributed. The curve will coincide with the x-axis if all the urban population is concentrated in one locality. The variation between hypothetical extremes of complete evenness and complete concentration is indicated by the degree to which it departs from the diagonal. A Lorenz curve is shown in Figure 8.3

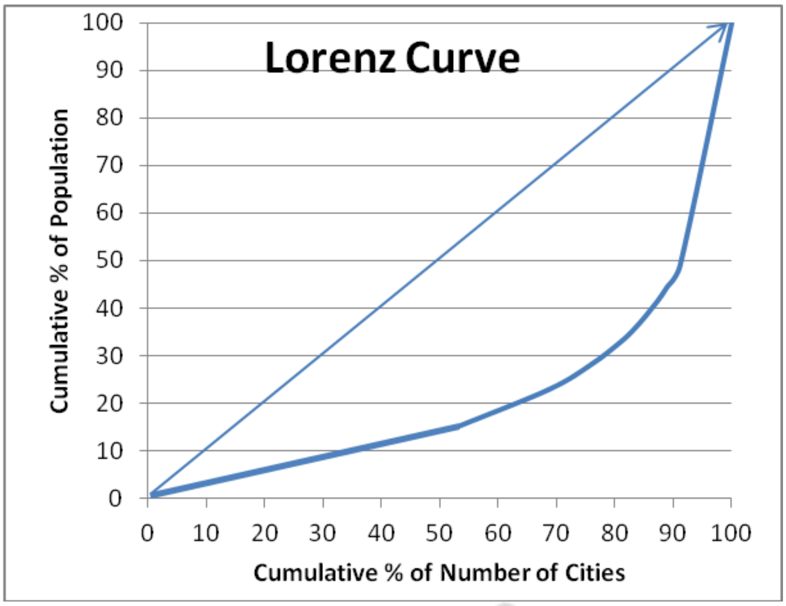

Figure 3 : Lorenz Curve
Source: Data above

1.4.3 Gini Concentration Index: This index measures the distribution of one variable relative to another. Gini concentration ratio measures the proportion of the total area under the diagonal and that of the area between the diagonal and the Lorenz curve. Thus, the area on the graph contained between the curve and the diagonal as a proportion of the entire area below the diagonal is expressed by the following

$$G_i = \sum_{i=1}^{n} X_i Y_{i+1} - \sum_{i=1}^{n} X_{i+1} Y_i$$

relationship :

where X_i and Y_i are the respective cumulative proportions of urban population and urban localities, and n is the number of class intervals or units. This index varies from zero when the population is evenly distributed and tends to unity when the population is concentrated in one unit. The Gini index can also be used to study the concentration of populations in cities. The higher the value of the index, the greater the level of concentration in the largest cities. Table 1 shows the Gini Concentration Ratios for India's urban population in 1991.

Table 1: Gini Concentration Ratio based on Urban Population of India, 1991.

Size Class	Urban Population ('000s)	Number of urban localities	Proportion of		Cumulative Proportion of	
			Urban Popu.	es	Urban Popu. X_i	Urban localities Y_i
5 million above	37,225	4	0.1749	0.0011	0.1749	0.0011
2 million - 5 million	16,261	5	0.0764	0.0014	0.2513	0.0025
1 million - 2 million	17,176	14	0.0807	0.0039	0.3320	0.0064
1 lakh - 1 million	68,141	273	0.3201	0.0756	0.6521	0.0820
50,000 - one lakh	23,309	341	0.1095	0.0945	0.7616	0.1765
20,000 - 50,000	28,079	927	0.1319	0.2568	0.8935	0.4333
10,000 - 20,000	16,531	1135	0.0776	0.3145	0.9711	0.7478
5,000 - 10,000	5,532	725	0.0260	0.2009	0.9971	0.9487
less than 5000	614	185	0.0029	0.0513	1.0000	1.0000
All classes	212,868	3609	1.0000	1.0000		

$= 3.0609 - 2.3405$

$= 0.7201$

1.4.4 Location Quotients: A location quotient can be used to compare one population characteristic with the total population in terms of its regional distribution. For example, the population of aged people (above 65 years) in a district A of state Z may be compared with the corresponding proportion in the total population of state Z. If the district has 36 per cent of the aged population, whereas the state as a whole has only 10 per cent of it, then obviously the aged population of the state is concentrated in district A. This concentration is expressed as a single statistic:

$$\frac{36}{10} = 3.6$$

Mathematically, it can be written as $LQ = \dfrac{x_i/p_i}{X/P}$

Where,
LQ is the Location Quotient,
x_i is the population of district i with characteristic x
p_i is the total population of district i
X is population of state having x characteristic
P is the total population of the state.

Location quotients below 1 express deficiency in a characteristic, and above 1 indicates its surplus.

1.5 Standard Distance

The spatial distribution of phenomenon can be studied based on the mean centre (also known as the centre of gravity) and standard distance, which is a measure of dispersion. Both mean centre and standard distance are like the arithmetic mean and standard deviation in statistics. Geographically, they are determined by the latitudes and longitudes of places, and weighted by population. Every individual is assumed to have equal weight and exerts an influence on a pivotal point, which is proportional to his distance from that point. A measure of dispersion around the centre of gravity is proposed as standard distance. On the other hand, the median centre is a somewhat different concept. It is a point of minimum aggregate travel. This is the point from which the sum of the linear radial deviations is a minimum. Therefore, it is a point of minimum aggregate travel. Median centre is an optimum location for accessibility of services and is an excellent tool in planning.

It should be noted that the location of the mean centre is affected by a change in the position of any unit of the population. It is most useful for studying the areal shifts of a distribution over time, whereas the median centre is useful for investigating locational optima for centralized services.

Further, concentration indices of urbanization do not take into consideration where the urban centres are located. Thus, these indices do not differentiate between countries which have their populous cities close together and those whose cities are dispersed. Population dispersion is a dimension of urbanization and reflects the distance between urban localities. This is exactly the opposite of concentration.

Standard Distance: This is a frequently used index for studying the dispersion of population. It can be interpreted as the standard deviation of the distribution of population from the centre of population. Let the country under study be divided into n localities. Let the co-ordinates of the locality i be x_i, y_i and the population living here be P_i. Let the total population of the country be P and the co-ordinates of the mean centre of population $\bar{x}$, $\bar{y}$.

$$SD = \sqrt{\frac{\sum_{i=1}^{n} P_i (x_i - \bar{x})^2 + \sum_{i=1}^{n} P_i (y_i - \bar{y})^2}{P}}$$

$$or \sqrt{\frac{\sum P_i K^2 (x_i - \bar{x})^2 + \sum P_i Q^2 (y_i - \bar{y})^2}{P}}$$

Then, the Standard distance is calculated by using the formula:

x_i = Horizontal co-ordinate of the locality i, which is degree of latitude of locality (i) × K.

y_i = Vertical co-ordinate of the locality i, which is degree of longitude of locality (i) x Q.

$$\bar{x} = \frac{\sum_{i=1}^{n} P_i x_i}{P} \qquad \bar{y} = \frac{\sum_{i=1}^{n} P_i y_i}{P}$$

where K and Q are constants for converting geographical degrees into linear units. Such kind of information is useful for infrastructural planning.

The higher the value of the index, the greater will be the degree of dispersion of localities. The standard distance has the same kind of relationship to the centre of population that standard deviation of a frequency distribution has with the arithmetic mean

References

- Duncan, O. D. (1957) "The Measurement of Population Distribution", *Population Studies*, Vol. 11, No. 1, pp. 27-45.

- Goldstein, S. and David F. Sly (eds), 1975, *The Measurement of Urbanization and Projection of Urban Population*, Chapter II, IUSSP, Ordina Editions, Belgium.

- Henry S. Shrylock, Jacob S. Siegal and Associates, 1980, *The Methods and Materials of Demography*, Vol II, US. Bureau of the Census, Washington D.C.

- United Nations, 1991, *World Urbanization Prospects*, 1990, New York.

- United Nations: *Determinants and Consequences of Population Growth and Trends*, New York, 1973.

- W.H.O (1988),"*Urbanization and its implications for Child Health"*, (WHO, Geneva).