

Religious Taxonomy of Indian States

THE distribution of population by religious affiliations has not received much attention from demographers in India even though data on religion on a national scale are available in the Census. The division of India in 1947 was the result of political rivalry based on religion. But these political differences could not have resulted in the creation of Pakistan if Muslims had not been concentrated so heavily in a few contiguous provinces. The distribution of Hindus and Muslims in pre-partitioned India clearly demarcated the sub-continent into two separate blocks (Mahalanobis, P. C., 1946). Since Independence, many new states have been formed and the boundaries of many have been redrawn. This too affected the religious composition of the states involved. It is of interest, therefore, to know the differences in the inter-state distribution of population according to religious affiliation and to investigate how the comparative position of States has changed over the past few decades.

Accordingly, the purpose of this paper is to investigate, (i) the inter-state differences as they existed in 1951, 1961 and 1971¹, (ii) the changes that have occurred in the religious composition of the states due to States Reorganisation, and (iii) whether the states that are similar in the religious distribution of population tend to form regions. The paper will also illustrate the application of certain statistical techniques to demographic problems.

1. A comparison of the differences in the states over the decade by keeping the areas constant is not of interest as low volume of internal migration could not have altered the religious composition considerably.

Method

The problem under investigation requires a pair-wise comparison of the population distribution by religion of the states and the grouping of the states that are similar in their composition. This requires, a measure to assess the difference between two distributions and a technique to classify states using this measure. In this paper, the distance function as defined by Bhattacharya, (Bhattacharya, A., 1946) will be used to compare population compositions and the algorithm of 'Hierarchical Clustering Scheme' (HCS) as proposed by Johnson (Johnson, S. C., 1967) will be used to group states.

The Distance Function

The concept of 'distance' defined for multinomial distributions by Bhattacharya is appropriate for problems of the present type. Matusita (Matusita, K., 1955, 1966) has shown that this measure can be used for decision purposes as it has desirable mathematical and statistical properties. The 'Bhattacharya distance' $d(X, Y)$ is defined as

$$d(X, Y) = [\sum(p_i^{1/2} - q_i^{1/2})^2]^{1/2}$$

where, the proportions p_i and q_i correspond to the two distributions X and Y . Though the statistic $d(\dots)$ is a measure of divergence, it can easily be shown to be functionally related to the measure of association called 'index of affinity' defined as

$$a(X, Y) = \sum(p_i q_i)^{1/2}$$

so that

$$d(X, Y) = [2\{1 - a(X, Y)\}]^{1/2}.$$

The statistic $a(\dots) = 1$, when $p_i = q_i$ for all $i = 1, 2, \dots, k$, i.e., when the two distributions are identical. And $a(\dots) = 0$ whenever for $p_i > 0$, $q_i = 0$ and vice-versa—a situation of complete divergence. The values of $d(\dots)$ corresponding to the extremes of $a(\dots)$ are 0 and $\sqrt{2}$. The distance function $d(\dots)$ satisfies the usual conditions of Euclidean metric. For any two distributions X and Y ,

$$d(X, Y) \geq 0; d(X, X) = 0$$

$$d(X, Y) = d(Y, X)$$

and

$$d(X, Y) + d(Y, Z) \geq d(X, Z).$$

Bhattacharya has shown the distribution of $d(\dots)$ as a function of chi-square

and examples of statistical tests using $d(\dots)$ can be had from Matusita's paper.

Hierarchical Clustering Scheme (HCS)

Once 'distance' between two units is defined, it is possible to group them on the criteria of shortest 'distance' or maximum 'affinity'. The method of clustering can be hierarchical or non-hierarchical. The former is essentially one of successively merging units that are at the 'shortest distance' (to be called 'group distance' here) to form a group/cluster. Thus at group distance of zero, there is likely to be as many groups as there are units. At greater group distances, units get merged and the number of groups get reduced, until finally at a certain group distance all the units merge to form a single group. Obviously hierarchical schemes tend to give unique classifications. For this reason HCS is adopted for this paper.

HCS will give different results depending on the way distance between groups is defined. In the process of forming groups, the distance $d[(x, y), (z)]$ from any newly formed group (x, y) to another unit/group (z) can be defined as the maximum (Johnson, S. C., 1967), minimum (Johnson, S. C., 1967), mean (Sneath, P. H. A. and Sokal, R. R., 1973)² or mean square (Chen, H. *et al*, 1974) of the distances $d(x, z)$ and $d(y, z)$. The HCS suggested by Johnson which considers

$$\max [d(x, z), d(y, z)] = d[(x, y), (z)]$$

is appropriate for this study. The 'maximum' tends to enlarge the number of small clusters and the 'minimum', to reduce the number of clusters. The minimum method has to be viewed as a grouping procedure whereby units that are close are linked one by one. This procedure is suitable when some sort of a 'tree type' connection exists between units. Such a relationship is not expected in the case of Indian states. The maximum method on the other hand encompasses units that are within a certain radius and as the radius increases more and more units get merged. The 'maximum' and 'minimum' methods are invariant under monotone transformations of the basic input information which will not be the case with the other methods mentioned above.

2. Median can also be used if there are at least three groups.

Data

The data on the distribution of population by religion for the States and Union Territories of India³ are taken from the Indian Census Reports of 1951, 1961 and 1971. The two decades witnessed considerable changes in the size and boundaries of states on account of reorganisation of States. However, for the problem formulated here, these changes do not necessitate adjustment of original data.

For the present study only Hindus, Muslims and Christians, who together form more than 95 per cent of the population, are considered separately. The rest of the groups mentioned in the census is treated as a single group and is called the "Rest". Census category of "others" is included in the "Rest". At the national level, the 'Rest' form a small group but in some states certain religious groups included therein constitute a majority, for example, the Sikhs in the Punjab and "others" in Arunachal Pradesh (in 1971). The Punjab and Arunachal Pradesh differ in their religious composition but these differences are not revealed in the present study because of this grouping of minority religious groups in the category of the 'Rest'.

Results

1. The differences in the inter-censal variations in population size of the religious groups of India between 1951 and 1971 have resulted in the reduction of the proportion of Hindus and an increase of the other groups (see Table 1).

The changes in religious composition of India's population between 1951 and 1971 are by no means drastic. The differences between the distributions measured by the 'distances' gave values 0.027 between 1951 and 1961, and 0.016 between 1961 and 1971. The distance between 1951 and 1971 worked out to be 0.036. The changes were greater during 1951-61 than during 1961-71. But the low values of the distance function reveal that the changes are insignificant.

3. For simplicity the term States will be used in the text hereafter to denote both States and Union Territories

TABLE 1—PERCENTAGE DISTRIBUTION OF THE VARIOUS RELIGIOUS GROUPS AND THEIR INTERCENSAL INCREASE RATE IN INDIA : 1951, 1961 AND 1971

<i>Year</i>	<i>Hindus</i>	<i>Muslims</i>	<i>Christians</i>	<i>The Rest</i>	<i>All-India</i>
1951	84.98	9.91	2.35	2.76	100.00
1961	83.51	10.70	2.44	3.35	100.00
1971	82.72	11.21	2.60	3.47	100.00
Rate of increase :					
1951-61	20.29	25.61	27.38	47.22	21.51
1961-71	23.69	30.85	32.60	27.45	24.80

2. With regard to comparative positions of the States the differences of the states from the Indian average are noted first. These distance values are given in Table 2.

For 1951, all the states are found to fall within a distance of 0.7. The re-organisation of the states during the 1950s appears to have had an averaging effect on the religious composition of the states. For, not only has the maximum distance of these states got reduced to 0.6, but many of the newly constituted states were at a lesser distance from the Indian average. Examples are Kerala and the Andamans.⁴ The exception is Bombay which in 1951, had a distribution closest to the Indian average. On its division, the religious composition of the two new states of Maharashtra and Gujarat became divergent. But the increase in the distance value *per se*, is only marginal.

The inclusion of Kashmir, Laccadives, etc., which had distributions very different from the Indian average has increased the maximum distance to 1.0, in 1961. But the distance got reduced to 0.9 in 1971, mainly due to changes in Nagaland and Laccadives. Figure 1 displays a greater dispersal of states in 1971 than in 1961.

4. For the sake of brevity only the first part of the names will be used in the case of States and Union Territories that are known by group names—e.g., Andaman and Nicobar Islands will be mentioned as 'Andamans' only. In the case of 'Jammu and Kashmir' however, we shall use 'Kashmir' to denote the state.

TABLE 2—DISTANCE OF THE STATES FROM INDIAN AVERAGE AS IN
1951, 1961 AND 1971

1951		1961		1971	
State	Dis.	State	Dis.	State	Dis.
1. Uttar Pradesh	0.159	Andhra Pradesh	0.173	Andhra Pradesh	0.172.
2. Bihar	0.056	Assam	0.209	Assam	0.212
3. Orissa	0.264	Bihar	0.077	Bihar	0.079
4. W. Bengal	0.166	Gujarat	0.108	Gujarat	0.123
5. Chandernagore	0.221	Jammu and Kashmir	0.641	Haryana	0.203
6. Assam	0.233	Kerala	0.383	Himachal Pradesh	0.262
7. Manipur	0.388	Madhya Pradesh	0.172	Jammu and Kashmir	0.608
8. Tripura	0.169	Madras	0.195	Kerala	0.383
9. Sikkim	0.483	Maharashtra	0.128	Madhya Pradesh	0.176
10. Madras	0.153	Mysore	0.098	Maharashtra	0.118
11. Mysore	0.122	Orissa	0.284	Manipur	0.398
12. Travancore- Cochin	0.465	Punjab	0.459	Meghalaya	0.823
13. Coorg	0.150	Rajasthan	0.147	Mysore	0.100
14. Bombay	0.053	Uttar Pradesh	0.170	Nagaland	0.948
15. Saurashtra	0.132	W. Bengal	0.180	Orissa	0.253
16. Kutch	0.265	Andaman and Nicobar Is.	0.437	Punjab	0.707
17. Madhya Pradesh	0.180	Delhi	0.138	Rajasthan	0.150
18. Madhya Bharat	0.145	Himachal Pradesh	0.256	Tamil Nadu	0.205
19. Hyderabad	0.114	Laccadive Minicoy and Amindivi Is.	1.070	Tripura	0.106
20. Vindhya Pradesh	0.250	Manipur	0.365	Uttar Pradesh	0.174
21. Bhopal	0.163	Tripura	0.143	West Bengal	0.175
22. Rajasthan	0.150	Dadra and Nagar Haveli	0.287	Andaman and Nicobar Is.	0.376
23. Punjab	0.449	Goa, Daman and Diu	0.520	Arunachal Pradesh	0.854
24. Himachal	0.268	Pondicherry	0.222	Chandigarh	0.397
25. PEPSU	0.622	NEFA	0.391	Dadra	0.268
26. Delhi	0.161	Nagaland	0.978	Delhi	0.146
27. Ajmer	0.114	Sikkim	0.454	Goa	0.459
28. Andaman	0.641	—		Laccadive	0.957
29. —		—		Pondicherry	0.224

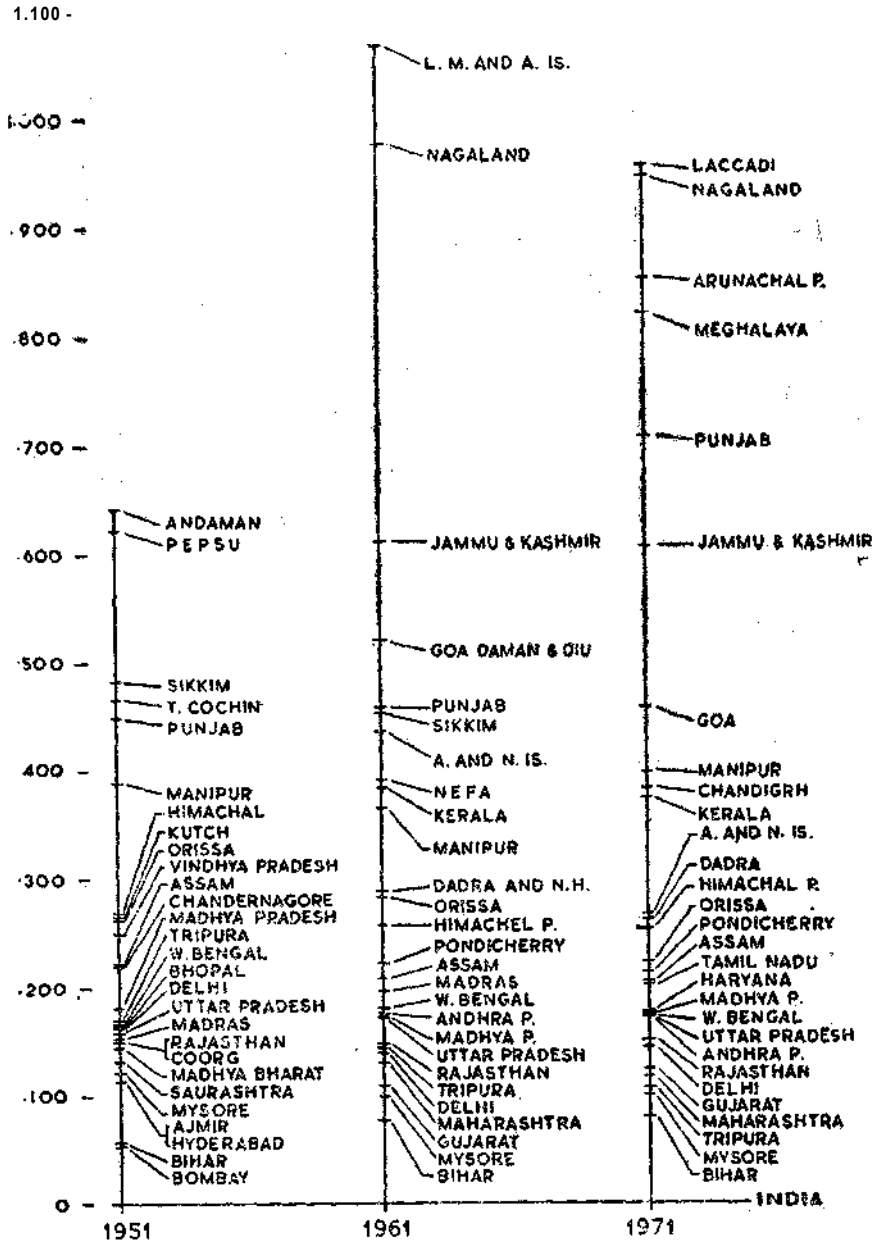


Fig. 1. Relative positions of the State of India as in 1951, 1961 and 1971 with respect to Indian average.

The formation of the state of Haryana from Punjab has resulted in the religious distribution of residual Punjab being much different from most of the states of India. At the same time Haryana has taken a place closer to Indian average. This kind of tendency can be noticed in the case of other states that were formed during the 1961-71 period.

3. The minimum distance at which all the states formed a single group as per the HCS followed here, increased from 0.86 in 1951 to 1.33 in 1961. The increase in distance should ordinarily be interpreted as greater diversification of the religious composition of the states. But the increase between 1951 and 1961 can be seen to be due to inclusion of Kashmir, Laccadives and Nagaland in 1961. If these three states are not considered, the others form a single group at a distance of 0.76—a value lower than in 1951. In other words, inter-state differences in religious composition have decreased during the decade. States Reorganisation, the major event of the decade that realigned administrative boundaries, can be said to have reduced religious diversity among the states.

Despite far reaching nature of the states reorganisation process of 1951-61, 16 of the 28 states of 1951 maintained their identity in 1961. The religious composition of these states had not changed significantly. Noticeable changes, however, are in Manipur (where proportion of Christians increased while that of 'Rest' decreased), Andamans (where proportion of Hindus increased and that of 'Rest' decreased) and Kerala (where proportion of Christians decreased and that of Muslims increased). The splitting of Bombay did not result in a greatly divergent composition for Gujarat and Maharashtra. The formation of Andhra Pradesh with the merging of parts of Madras state with Hyderabad did not produce divergent distributions.

Of the 10 States of 1951 that merged with the neighbouring states only 3, namely, Chandranagore, Kutch and Pepsu, had distributions different from the bigger states with which they merged. But these differences have not altered the percentage distribution of the composite states.

Seven new states came into being by 1961. Of these Kashmir and Laccadives were dominated by Muslims. Nagaland had Christians and 'Rest' in large proportions and the other states were Hindu dominated.

The changes described above explains the reduction of inter-state differences in 1961. The inclusion of Muslim states and a small increase in the proportion of Muslims in some of the states have accentuated inter-states dissimilarities only marginally.

In 1971 the minimum distance value decreased to 1.25 from 1.33 in 1961. A further look at the HCS groups reveals that distances were reduced much more than is indicated by this small decrease in the minimum distance value. The formation of new states by the division of Punjab and the regrouping of the North-Eastern region during 1961-71 have, however, created groups (Punjab, Arunachal Pradesh and Meghalaya) that are at a greater distance than the earlier states were in 1961. For the remaining states of 1951 (excluding Naga-land, Kashmir and Laccadives) the distance at which they form a single group in 1971 is found to be 0.44, which is much lower than 0.86 and 0.76 in 1951 and 1961. The groups formation in 1971 shows that Nagaland merged with other States at a distance of 1.33 in 1961. The group consisting of Kashmir and Laccadives, however, has not reduced the inter-state distance significantly. The reduction of distance for Laccadives to only by 0.03 units—from 1.28 in 1961 to 1.25 in 1971.

The territorial changes of 1961-71 left twenty three states unaffected. From among these, 23 states, substantial changes in the religious composition were noticeable in the cases of Tripura, Manipur, Nagaland and Andamans. The proportion of Hindus increased significantly in Tripura; and Andamans. In Manipur and Nagaland it was a case of a reduction in the percentage of others compensated by an increase of Christians.

The division of Punjab has resulted in the new states being divergent in their distributions. Haryana has become a predominantly Hindu states while the concentration of Sikhs increased in the residual Punjab. Similarly, the reorganisation of the North-eastern region created states like Arunachal Pradesh and Meghalaya, where the unspecified "others" were in majority in the former and of Christians and others in the latter.

The lower growth rate among the Hindus has reduced the proportion of Hindus in many states. Reorganisation of the states during 1961-71 has increased the number of states where Christians and 'The Rest' have become the largest groups. These changes produced many outliers and thus accounts for

the larger number of smaller groups in 1971 than in 1961. Hence though inter-state distances decreased in many cases especially among the Hindu majority states, inter-group distances had increased.

In the previous section we have seen that inter-state distances have decreased over the years in many cases. The tendency was for states that were close, to move even closer. This tended to widen inter-group differences and increased the number of clusters. This is evident from Table 3 given below.

TABLE 3-NUMBER OF HCS GROUPS AT VARIOUS GROUP DISTANCES

<i>Group distance</i>	<i>Number of groups</i>		
	<i>1951</i>	<i>1961</i>	<i>1971</i>
0	28	27	29
Less than 0.10	17	19	20
0.11-0.20	11	12	15
0.21-0.30	7	10	10
0.31-0.40	5	7	8
0.41-0.50	4	6	6
0.51-0.60	3	5	4
0.61-0.70	2	4	4
0.71-0.80	2	3	4
0.81-0.90	1	3	3
0.91-1.00	1	3	3
1.01-1.10	1	3	2
1.11-1.20	1	2	2
1.21-1.30	1	2	1
1.31-1.40	1	1	1

The table shows the number of groups at each group distance. At distance of 0.1, for instance, the 28 states of 1951 formed themselves into 17 groups, the 27 states of 1961 into 19 groups and the 29 states of 1971 into 20 groups. At a distance of about 0.5, where the grouping is more or less clear cut, there are more groups in 1961 and 1971 than in 1951. At distances greater than

0.5, the groups do not appear to merge faster. In fact, if the religious composition of the states of the different groups are also taken into account, the distinctness of the groups can be easily explained to be due to the decrease in the percentage of Hindus in many of the states that are in the extreme groups.

5. The changes in the religious composition of a few states especially in the newly formed ones has been suggested above as the reason for the increase in the number of groups. Hindus form the biggest religious group in almost all states—some with more than 95 per cent even. States where Hindus are less than 50 per cent are few. They numbered two (Andamans (30 per cent Hindus) and Pepsu (49 per cent) in 1951; three (Kashmir (29 per cent), Laccadives (1 per cent) and Nagaland (9 per cent) in 1961; and six (Punjab (38 per cent), Arunachal Pradesh (22 per cent), Nagaland (11 per cent), Meghalaya (18 per cent), Kashmir (30 per cent), and Laccadives (5 per cent)) in 1971.

The changes in the membership of the groups over the years are not strictly comparable because of territorial changes. But an overall picture can be delineated keeping in mind only the broad geographical locations. It can be observed that changes are very minimal. The present states of Orissa, Madhya Pradesh and Himachal Pradesh have consistently remained as a group over the years. In all these three states more than 95 per cent of the population is Hindu. The states of Maharashtra, Delhi, Gujarat, Rajasthan, Tripura, Bihar, Mysore, Uttar Pradesh, West Bengal, Andhra Pradesh and Tamil Nadu, which again are Hindu dominated states, can be identified as the second major group. The changes in this group are also minimal. Significant ones, however, are the addition of Haryana and Assam in 1971. In this group of states Andhra Pradesh and Tamil Nadu have identified themselves consistently as a sub-group, which had a position very close to the Hindu concentrated states in 1951 but has since experienced a shift in position. Mysore, a southern state has a religious distribution different from that of Andhra Pradesh and Tamil Nadu, but closer to that of Bihar and Uttar Pradesh. Another southern state, Kerala, is grouped with Andamans, Manipur and Goa where Christians are in good proportion. Assam which was in the Christian group in 1961 has moved into the Hindu group in 1971. But the new state of Arunachal Pradesh formed in the Eastern region, gets grouped with Punjab and Chandigarh⁵, while the new

5. It may be reminded here that these states form a group by themselves due to closeness of their religious distributions, they differ when the religions under 'the rest' are identified separately.

state of Meghalaya forms a group with Nagaland as the other member. The group of States mentioned last can be considered as 'the Rest' dominated one. But the extreme group consisting of Kashmir and Laccadives maintained identity throughout without additions or deletions.

Though group membership of the two major Hindu clusters has not changed significantly, the new groups formed out of territorial changes consist of states that have religious distributions different from the majority of the states of India. In 1971 there are 4 groups with 11 states where Muslims, Christians and 'The Rest', or combinations of them are concentrated, while there were only 2 non-Hindu groups in 1951 with 6 states. The inclusion of some of the large states in these non-Hindu groups makes them important in the study of social and political pressures in India. In this context the formation of more sub-groups within the Hindu dominated groups also deserves attention. The most significant aspect in the group formation is that the states within the groups are not contiguous in all cases.

Conclusion

Though the religious distribution in India as a whole has not changed considerably over the past two decades, inter-state differences seem to have increased as a combined result of the reorganisation of the states and the religious differentials in rate of population growth within the states. The Hindus remained the predominant religious group in almost all states of India in 1951. By 1971, many states have emerged with the other religious groups in larger proportions. In spite of the inter-state differences in religious composition decreasing in the case of a few states especially in the Hindu dominated areas of 1951, there was greater diversification as was evident from the larger number of prominent groups. Not only that, the group characteristic has become more specific so as to classify them as being dominated by a particular religion. In other words, a certain amount of polarisation has been in operation due to the decrease of inter-state distances within groups accompanied by a widening of the inter-group differences. But the important point is that states within a group are not always contiguous.

The emergence of groups of states with divergent religious distributions deserve attention. It raises several pertinent questions particularly about national integration but these questions cannot be answered without detailed investigation about the social and political behaviour in the groups of States.

For instance, the voting behaviour in these groups of states can bring out the influence the religion has on the electorate.

References

1. Bhattacharya, A., 1946, On a measure of divergence between two multinomial populations, *Sankhya*, 7(4), 401-406.
2. Chen, H., Ganadesikan, R. and Kettenring, J. R., 1974, Statistical methods for grouping corporations, *Sankhya*, 36, Series B, Part 1.
3. Johnson, S. C., 1967, Hierarchical clustering scheme, *Psychometrika*, 32(3).
4. Mahalanobis, P. C., 1946, Distribution of muslims in the population of India, *Sankhya*, 7(4).
5. Matusita, K., 1955, Decision rules, based on the distance for problems of fit, two samples and estimation, *Annals of Mathematical Statistics*, 26.
6. .____, 1966, A distance and related statistics in multivariate analysis. In : P.R. Krishnaiah (ed.), *Multivariate Analysis*, Academic Press, New York.
7. Sneath, P. H. A. and Sokal, R. R., 1973, *Numerical Taxonomy*, Freeman and Co., San Francisco,