

ANALYSIS OF VARIANCE OF DEMOGRAPHIC VARIABLES

By RANJAN KUMAR SOM

and

HIRALAL MUKHERJEE

Indian Statistical Institute

SUMMARY. The arrangement of the total sample in the form of interpenetrating net-works of sub-samples enable one to test by the method of analysis of variance the differential effects of the investigators (parties) and of other factors causing variation. Mahalanobis (1958) made the analysis of variance of four demographic variables, namely, the sex-ratio, the percentages of population in labour force, population literate, and population aged 50 years and above on the basis of the ninth round (May-November 1955) of the National Sample Survey, taking the estimates at the State level as the basic observation. In this note, the study has been extended by taking the estimates at the village level as the basic observation for the State of Uttar Pradesh. The same general conclusion emerges, namely, that there was no significant differential investigator effect, that is, that the survey was generally under statistical control.

1. INTRODUCTION

P. C. Mahalanobis (1958) had presented in a lecture series in Japan in 1958 the results of analysis of variance of four demographic variables, namely, sex-ratio (number of males per 100 females), percentage of population in labour force, percentage of literate population, and percentage of population aged 50 years and above. For each of the 19 States or State-groups, with two parties of investigators working in two time periods, the four party \times time combinations were taken as the basic observation; and the analysis of variance was done within each State with the following partition:

	d.f.
time	1
party	1
error	1 (party \times time)
<hr/>	
total	3

For each of the characteristics, out of the 19 F -ratios for each of Party/Error and Time/Error, none or only one came out significant at 5 per cent level, which was as could be expected. The study has been extended in this note; the data for the State of Uttar Pradesh (rural) being analyzed separately for strata having 4, 8 and 12 sample villages each; the data for each village being taken as the basic observation.

2. SAMPLE DESIGN AND PROCEDURE

The design was two-stage stratified. Rural India was divided into 256 ultimate strata consisting of districts or groups of adjoining districts, and within each stratum the sample villages were allocated in proportion to its 1951 Census population; then the allocated numbers were adjusted to make them 4 or multiples of 4. That is, in other words, within each of the strata the sample villages, which were 4 (or multiples of 4) in numbers, were selected with probability proportional to 1951 Census population and with replacement; in each sample village thus selected, 10 households were selected systematically with a random start. The survey was so designed that in each State there were two independent "parties" of investigators, each working in two different time periods, consisting of three months each—mid-May to mid-August and mid-August to mid-November 1955. In Uttar Pradesh,

SANKHYĀ : THE INDIAN JOURNAL OF STATISTICS : SERIES B

280 sample villages were selected from 46 strata—25 strata with 4 sample villages each, 18 strata with 8 sample villages each, and 3 strata with 12 sample villages each.

The arrangement of the total sample in the form of interpenetrating networks of sub-samples enables one to test by the method of Fisherian analysis of variance the effect of the investigators (parties) and of other factors causing variation (Mahalanobis, 1946). It is easy to write the mathematical models involved. W. G. Cochran (1953) has, for example, given two such models with parties and replications.

For the present study of analysis of variance, the characteristic for each sample village was taken as the basic observation. The analysis of variance was done with each of the 46 strata into which the State was divided with the following three different partitions:

(A) <i>Within stratum having</i> 4 sample villages each		(B) <i>Within stratum having</i> 8 sample villages each		(C) <i>Within stratum having</i> 12 sample villages each	
	d.f.		d.f.		d.f.
party	1	party	1	party	1
time	1	time	1	time	1
error	1 (party × time)	party × time	1	party × time	1
		error	4	error	8
<hr/>		<hr/>		<hr/>	
total	3	total	7	total	11
<hr/>		<hr/>		<hr/>	

The universe was not normal, and strictly speaking Fisher's F -ratio cannot be used to test the significance. The use of F may however, as made by Mahalanobis (1946, 1958), give approximate results.

3. RESULTS

Tables 1.1, 1.2, 1.3 and 1.4 give the analysis for sex-ratio, percentage of population in labour force, percentage of literate persons and percentage of population aged 50-above respectively for each of the stratum having 4 sample villages each; Tables 2.1, 2.2, 2.3 and 2.4 corresponding analysis for each of the stratum having 8 sample villages each; and Tables 3.1, 3.2, 3.3 and 3.4 that for each of the stratum having 12 sample villages each.

Table 4 shows the number of F -ratios significant at 5% level (as also those significant at 1% level) for each of the four demographic characteristics. Table 4 also shows for strata having 8 and 12 sample villages each, the number of significant (at 5% and 1% levels) F -ratios for party/(party × time) and time/(party × time): as the true "error" mean squares are available, these are shown only for illustrative purposes.

For the F -ratios time/(party × time) in strata with 4 sample villages each, out of the total 25 for each characteristic, a maximum of 3 significant F -ratios for sex-ratio among the four demographic characteristics is observed at 5% level. This may probably be due to the differential migration of the rural population in the two time periods. For the F -ratios time/error in the 18 strata with 8 sample villages each, one each has come out significant for the three characteristics—sex-ratio, percentage in labour force, percentage literate.

ANALYSIS OF VARIANCE OF DEMOGRAPHIC VARIABLES

TABLE 1.1. ANALYSIS OF VARIANCE OF SEX-RATIO (100 MALES/FEMALES) BY STRATUM WITH FOUR VILLAGES

serial no.	stratum no.	sum of squares				F-ratio	
		total	party	time	error	party	time
(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	02	2901.90	373.65	664.81	1403.64	0.20	0.30
2	10	2274.47	2079.82	157.38	37.27	85.80	4.22
3	11	816.42	641.86	173.84	0.72	891.47*	241.44*
4	14	725.35	450.69	85.10	189.75	2.37	0.45
5	16	2879.68	50.77	1767.79	1061.12	0.05	1.87
6	17	1713.12	447.11	455.18	810.83	0.65	0.56
7	18	841.39	713.96	45.16	82.27	8.68	0.53
8	23	295.71	168.99	108.99	81.73	1.33	1.33
9	24	1029.13	57.08	707.83	264.22	0.22	2.68
10	25	728.17	22.85	684.35	20.97	1.09	32.63
11	26	357.92	23.71	280.23	51.98	0.40	5.39
12	27	5490.28	3503.46	1611.22	376.68	9.33	4.29
13	29	1530.73	484.88	40.90	1004.89	0.48	0.04
14	33	1319.17	719.05	418.82	181.30	3.97	2.31
15	34	1356.46	865.35	738.75	112.36	4.50	6.87
16	35	4304.03	904.72	400.23	2930.06	0.33	0.14
17	36	831.41	458.89	213.24	561.28	0.08	0.30
18	39	759.09	132.83	665.90	1.17	112.63	569.15*
19	40	2778.32	1790.06	801.23	381.03	4.71	1.58
20	41	9149.00	2738.05	4138.99	2421.00	1.21	1.82
21	42	399.87	358.08	3.31	40.58	8.77	0.08
22	43	2395.65	310.11	2083.01	2.53	122.57	823.32*
23	44	2128.74	1674.44	16.64	437.66	3.83	0.04
24	45	827.85	51.05	686.17	90.73	0.56	7.58
25	46	233.78	148.41	70.60	16.81	8.71	4.20

*Significant at 5% level: $F_{5\%}(1,1) = 161.4$, $F_{1\%}(1,1) = 4032$.

TABLE 1.2. ANALYSIS OF VARIANCE OF PERCENTAGE OF POPULATION IN LABOUR FORCE BY STRATUM WITH FOUR VILLAGES

serial no.	stratum no.	sum of squares				F-ratio	
		total	party	time	error	party	time
(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	02	544.00	1.08	41.03	501.98	0.00	0.08
2	10	389.44	344.20	8.73	36.42	9.45	0.24
3	11	110.82	26.01	86.12	4.70	5.41	17.98
4	14	163.70	49.07	40.64	74.05	0.66	0.66
5	16	69.56	17.47	2.62	49.67	0.33	0.05
6	17	89.02	6.40	0.35	82.27	0.08	0.00
7	18	78.89	41.47	36.72	0.60	82.94	73.44
8	23	874.67	611.21	100.00	63.36	9.07	1.63
9	24	398.04	170.56	108.01	120.87	1.41	0.89
10	25	91.62	4.22	72.65	14.55	0.29	5.01
11	26	67.45	1.22	25.25	70.98	0.02	0.36
12	27	66.84	38.32	6.67	22.65	1.69	0.25
13	29	28.47	17.43	6.74	4.30	4.03	1.57
14	33	177.62	6.97	169.26	1.39	5.01	121.77
15	34	248.15	189.47	2.35	66.33	3.36	0.04
16	35	253.42	0.01	29.76	223.65	0.00	0.13
17	36	26.15	1.59	0.68	24.60	0.06	0.00
18	39	219.25	49.07	0.91	170.17	0.29	0.00
19	40	371.03	134.67	64.24	182.12	0.74	0.30
20	41	107.08	62.27	6.95	48.86	1.07	0.12
21	42	216.47	75.87	82.00	58.51	1.30	1.40
22	43	257.85	18.28	102.11	137.47	0.13	0.74
23	44	66.06	6.04	30.64	29.38	0.25	1.60
24	45	323.36	200.39	11.77	102.20	2.06	0.12
25	46	484.01	9.12	82.43	392.44	0.02	0.21

 $F_{5\%}(1,1) = 161.4$, $F_{1\%}(1,1) = 4032$.

SANKHYĀ : THE INDIAN JOURNAL OF STATISTICS : SERIES B

TABLE 1.3. ANALYSIS OF VARIANCE OF PERCENTAGE OF LITERATE PERSONS BY STRATUM WITH FOUR VILLAGES

serial no.	stratum no.	sum of squares				F-ratio	
		total	party	time	error	party	time
(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	02	296.09	121.88	3.04	75.17	1.73	0.04
2	10	305.69	12.82	65.77	227.10	0.06	0.29
3	11	157.13	42.96	25.15	89.02	0.48	0.28
4	14	160.98	0.01	8.33	152.61	0.00	0.05
5	16	18.31	1.28	14.59	2.44	0.62	3.98
6	17	605.62	1.00	591.04	12.68	0.08	46.64
7	18	114.64	13.80	23.57	77.17	0.18	0.31
8	23	56.07	1.51	14.74	39.82	0.04	0.37
9	24	37.12	26.93	0.32	9.87	2.73	0.03
10	25	95.16	22.51	69.07	2.68	8.40	26.11
11	26	47.68	6.83	15.80	24.05	0.27	0.63
12	27	63.02	26.09	22.94	1.09	14.12	11.63
13	29	296.62	202.02	92.64	1.06	101.43*	87.40
14	33	62.17	14.03	45.10	3.04	4.02	14.84
15	34	160.50	66.91	81.36	12.32	5.43	6.60
16	35	679.32	85.10	421.69	172.53	0.49	2.44
17	36	34.06	10.06	14.52	0.48	1.16	1.63
18	39	152.58	0.00	154.01	36.57	0.00	3.99
19	40	161.57	64.16	1.17	36.24	1.77	0.03
20	41	470.67	45.97	0.98	432.64	0.11	0.60
21	42	208.18	140.07	66.67	1.42	98.64	46.05
22	43	594.66	43.04	173.71	377.91	0.11	0.46
23	44	222.59	166.28	55.88	0.43	396.70*	120.05
24	45	625.27	478.30	17.81	29.16	16.40	0.61
25	46	531.03	137.13	79.04	315.76	0.43	0.25

*Significant at 5% level: $F 5\%(1, 1) = 161.4$, $F 1\%(1, 1) = 4052$.

TABLE 1.4. ANALYSIS OF VARIANCE OF PERCENTAGE OF POPULATION AGED 50-ABOVE BY STRATUM WITH FOUR VILLAGES

serial no.	stratum no.	sum of squares				F-ratio	
		total	party	time	error	party	time
(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	2	11.71	3.33	7.64	1.34	2.49	6.25
2	10	100.70	66.68	11.00	3.02	28.70	2.07
3	11	11.22	1.67	0.45	0.10	16.70	94.50
4	14	134.10	49.09	21.16	63.04	0.79	0.34
5	16	64.00	1.08	17.22	45.70	0.02	0.38
6	17	13.82	0.27	0.20	13.35	0.02	0.01
7	18	47.86	16.61	5.09	26.16	0.63	0.19
8	23	61.43	2.56	14.29	34.58	0.07	0.41
9	24	92.07	1.07	38.28	53.22	0.62	0.72
10	25	8.65	2.58	4.75	0.62	5.29	7.66
11	26	80.00	0.03	35.22	63.81	0.00	0.65
12	27	203.24	105.30	5.74	2.20	88.77	2.61
13	29	83.15	27.25	16.08	39.82	0.68	0.40
14	33	174.32	37.03	88.64	48.65	0.76	1.82
15	34	172.00	22.00	67.01	63.80	0.24	0.61
16	35	102.11	24.46	76.48	1.17	20.01	65.37
17	36	73.30	11.19	66.03	0.17	1.81	0.88
18	39	20.20	28.09	0.60	0.64	43.80	0.88
19	40	72.85	4.21	38.94	29.70	0.14	1.21
20	41	22.63	10.81	5.38	0.60	25.47	8.15
21	42	108.52	81.00	4.25	22.37	3.00	0.10
22	43	62.43	0.01	0.24	62.18	0.00	0.00
23	44	49.77	2.73	37.14	0.99	0.28	3.75
24	45	73.94	62.06	0.13	21.75	2.39	0.01
25	46	153.63	63.07	18.10	82.36	0.64	0.22

 $F 5\%(1, 1) = 161.4$; $F 1\%(1, 1) = 4052$.

ANALYSIS OF VARIANCE OF DEMOGRAPHIC VARIABLES

TABLE 2.1. ANALYSIS OF VARIANCE OF SEX-RATIO (100 MALES/FEMALES) BY STRATUM WITH EIGHT VILLAGES

serial no.	stratum no.	sum of squares					mean square for error	F-ratio		
		total	party	time	party x time	error		party	time	party x time
(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	01	5222.71	12.52	824.00	848.10	3737.10	931.28	0.01	0.67	0.91
2	05	8403.81	1097.46	1407.13	4584.99	1374.21	343.55	3.19	4.10	13.35
3	06	3035.77	710.01	1471.80	146.81	697.25	174.31	4.13	8.44	0.84
4	07	1609.52	0.02	8.67	981.04	679.79	169.05	0.00	0.05	0.77
5	08	3122.31	1315.33	111.90	674.92	1020.10	255.04	5.10	0.44	2.65
6	09	2532.98	204.03	73.09	1343.51	908.55	227.14	0.90	0.33	5.92
7	13	7100.04	4.31	733.33	3328.00	3044.34	761.08	0.01	0.05	4.27
8	15	12921.79	745.36	4001.54	824.47	7550.42	1887.60	0.39	2.12	0.33
9	19	2802.40	158.80	175.87	2008.20	459.53	114.88	1.28	1.53	17.48
10	20	4292.08	100.32	742.47	609.04	2750.25	687.56	0.15	1.08	1.02
11	21	10130.74	177.19	261.02	3258.25	6230.27	1280.07	0.11	0.24	2.06
12	22	2878.22	24.10	373.61	65.26	2415.16	603.79	0.04	0.62	0.11
13	28	5470.28	173.78	1644.51	334.12	3321.87	830.47	0.21	1.04	0.40
14	30	2314.43	109.00	81.08	141.88	1890.67	472.07	0.42	0.17	0.30
15	31	5713.49	0.29	1830.91	1.76	3872.53	968.13	0.00	1.90	0.00
16	32	1525.15	190.90	89.73	57.79	1206.73	301.68	0.63	0.23	0.19
17	37	4784.12	139.36	611.62	726.38	3284.76	821.69	0.17	0.74	0.88
18	38	2456.03	840.29	43.47	5.57	1567.30	391.82	2.14	0.11	0.01

*Significant at 5% level: $F 5\%(1, 4) = 7.7$; $F 1\%(1, 4) = 21.2$.

TABLE 2.2. ANALYSIS OF VARIANCE OF PERCENTAGE OF POPULATION IN LABOUR FORCE BY STRATUM WITH EIGHT VILLAGES

serial no.	stratum no.	sum of squares					mean square for error	F-ratio		
		total	party	time	party x time	error		party	time	party x time
(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	01	1143.18	137.28	14.00	169.46	821.75	205.44	0.67	0.07	0.82
2	05	335.21	27.00	12.44	4.27	300.90	75.22	0.37	0.03	0.08
3	06	643.45	22.47	3.55	241.46	375.97	93.99	0.24	0.04	2.57
4	07	228.55	66.47	26.67	24.02	162.59	25.65	2.59	1.04	0.97
5	08	345.19	7.22	115.06	188.19	34.81	8.65	0.85	13.30*	21.76**
6	09	1622.14	6.23	0.23	410.41	1205.27	301.32	0.02	0.90	1.50
7	13	871.63	0.85	432.62	0.01	438.16	109.54	0.01	3.95	0.00
8	15	1067.72	33.66	200.20	7.91	785.89	196.47	0.17	1.32	0.04
9	19	397.37	97.65	5.67	12.47	281.98	70.42	1.39	0.08	0.18
10	20	693.08	173.26	4.10	162.60	413.06	103.26	1.68	0.04	0.99
11	21	595.59	472.78	8.04	84.81	65.00	16.40	26.67**	0.49	2.86
12	22	369.68	121.84	38.02	0.50	200.32	52.33	2.33	0.73	0.01
13	28	207.43	27.00	22.56	86.37	71.60	17.88	1.51	1.26	4.83
14	30	756.82	191.78	0.04	313.39	251.01	62.90	3.05	0.60	4.98
15	31	147.80	0.06	15.20	113.79	18.75	4.69	0.01	3.24	21.26**
16	32	149.76	41.45	16.16	13.01	78.24	19.56	2.12	0.83	0.71
17	37	235.25	12.78	31.32	91.46	96.69	24.17	0.53	1.42	3.78
18	38	720.66	65.49	41.36	354.18	250.63	64.01	1.01	0.64	3.40

*Significant at 5% level: $F 5\%(1, 4) = 7.7$.**Significant at 1% level: $F 1\%(1, 4) = 21.2$.

SANKHYĀ : THE INDIAN JOURNAL OF STATISTICS : SERIES B

TABLE 2.3. ANALYSIS OF VARIANCE OF PERCENTAGE OF LITERATE PERSONS BY STRATUM WITH EIGHT VILLAGES

serial no.	stratum no.	sum of squares					mean square for error	F-ratio		
		total	party	time	party x time	error		party	time	party x time
(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	01	2238.21	773.03	203.81	614.25	587.12	146.78	5.27	1.80	4.18
2	05	506.20	330.50	6.23	24.12	246.26	61.56	3.75	0.08	0.39
3	06	282.42	63.22	18.37	2.24	198.39	49.60	1.27	0.37	0.95
4	07	397.13	3.37	82.61	07.00	214.05	53.51	0.06	1.54	1.81
5	08	800.18	17.88	35.11	9.29	737.00	184.48	0.10	0.19	0.05
6	09	807.18	5.52	1.72	78.86	721.08	180.27	0.03	0.01	0.44
7	13	751.63	121.01	10.33	68.88	548.51	130.63	0.89	0.12	0.49
8	15	1017.07	45.60	710.34	2.67	249.46	62.36	0.73	11.64	0.40
9	19	1025.01	33.78	4.47	217.90	769.67	192.42	0.18	0.02	1.13
10	20	502.04	0.50	16.70	44.37	531.31	132.83	0.00	0.13	0.33
11	21	240.16	4.02	45.74	17.07	181.43	45.36	0.69	1.01	0.40
12	22	224.76	68.64	30.20	5.95	130.01	32.50	1.80	0.93	0.18
13	28	504.41	0.02	10.75	130.70	347.04	86.76	0.01	0.23	1.68
14	30	371.28	84.24	74.11	28.17	204.78	51.19	1.25	1.45	0.65
15	31	396.90	8.72	108.42	95.28	184.48	46.12	0.19	2.35	2.07
16	32	609.75	104.98	120.61	129.76	355.40	83.85	1.25	1.53	1.53
17	37	335.34	101.24	17.40	01.00	125.70	31.42	3.22	0.55	2.90
18	38	2459.12	450.90	490.47	175.78	1341.07	335.49	1.34	1.46	0.52

*Significant at 5% level: $F 5\%(1, 4) = 7.7$; $F 1\%(1, 4) = 21.2$.

TABLE 2.4. ANALYSIS OF VARIANCE OF PERCENTAGE OF POPULATION AGED 30 ABOVE BY STRATUM WITH EIGHT VILLAGES

serial no.	stratum no.	sum of squares					mean square for error	F-ratio		
		total	party	time	party x time	error		party	time	party x time
(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	01	102.27	5.12	2.06	33.05	152.04	38.01	0.13	0.65	0.87
2	05	353.00	12.75	62.72	95.64	181.89	45.47	0.28	1.38	2.10
3	06	181.28	75.46	14.24	10.42	81.16	20.19	3.72	0.70	0.51
4	07	144.80	7.20	16.10	3.97	117.53	29.38	0.25	0.55	0.14
5	08	111.04	19.72	0.00	21.26	70.06	17.52	1.13	0.90	1.21
6	09	86.62	3.14	16.66	12.37	56.55	14.14	0.22	1.17	0.87
7	13	205.70	44.61	5.63	0.19	155.27	38.82	1.15	0.14	0.09
8	15	303.43	61.27	0.82	64.07	177.27	44.32	1.38	0.02	1.45
9	19	131.51	39.78	57.67	0.14	36.92	9.23	4.31	6.25	0.62
10	20	289.97	0.26	17.61	3.63	268.47	67.12	0.00	0.26	0.05
11	21	82.78	0.73	2.06	29.11	50.88	12.72	0.06	0.16	2.29
12	22	413.89	1.24	22.12	47.03	343.50	83.88	0.01	0.26	0.55
13	28	265.53	26.01	0.04	22.68	215.30	53.82	0.40	0.02	0.42
14	30	140.16	2.10	3.66	38.82	95.62	23.00	0.09	0.15	1.62
15	31	169.43	14.03	62.61	8.08	83.70	20.65	0.67	2.00	0.43
16	32	89.80	37.84	0.08	35.62	16.43	4.11	0.21	0.02	8.07
17	37	282.68	2.50	84.05	40.18	155.95	38.99	0.06	2.16	1.03
18	38	161.58	10.03	4.25	94.32	62.08	13.02	0.84	0.33	7.24

*Significant at 5% level: $F 5\%(1, 4) = 7.7$; $F 1\%(1, 4) = 21.2$.

ANALYSIS OF VARIANCE OF DEMOGRAPHIC VARIABLES

TABLE 3.1. ANALYSIS OF VARIANCE OF SEX-RATIO (100 MALES/FEMALES) BY STRATUM WITH TWELVE VILLAGES

serial no.	stratum no.	sum of squares					mean square for error	F-ratio		
		total	party	time	party X time	error		party	time	party X time
(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	03	9301.78	40.89	1274.11	156.32	7830.44	978.81	0.04	1.30	0.16
2	04	11410.66	3604.93	91.57	9.99	7704.17	963.02	3.74	0.10	0.01
3	12	4933.24	31.83	568.42	92.68	4240.31	530.04	0.06	1.07	0.17

$F 5\%(1, 8) = 5.32$; $F 1\%(1, 8) = 11.26$.

TABLE 3.2. ANALYSIS OF VARIANCE OF PERCENTAGE OF POPULATION IN LABOUR FORCE BY STRATUM WITH TWELVE VILLAGES

serial no.	stratum no.	sum of squares					mean square for error	F-ratio		
		total	party	time	party X time	error		party	time	party X time
(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	03	1776.20	326.86	366.97	343.61	738.76	92.34	3.64	3.97	3.72
2	04	2109.63	16.05	568.84	0.82	1514.92	189.36	0.08	3.00	0.05
3	12	376.80	28.93	25.58	63.28	539.01	32.38	0.89	0.77	1.95

$F 5\%(1, 8) = 5.32$; $F 1\%(1, 8) = 11.26$.

TABLE 3.3. ANALYSIS OF VARIANCE OF PERCENTAGE OF LITERATE PERSONS BY STRATUM WITH TWELVE VILLAGES

serial no.	stratum no.	sum of squares					mean square for error	F-ratio		
		total	party	time	party X time	error		party	time	party X time
(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	03	821.94	72.90	25.05	103.48	818.51	77.31	0.94	0.32	1.36
2	04	888.79	20.53	107.88	96.56	603.82	82.98	0.25	1.30	1.16
3	12	1425.80	10.86	168.45	56.13	1190.56	170.04	0.06	0.99	0.33

$F 5\%(1, 8) = 5.32$; $F 1\%(1, 8) = 11.26$.

SANKHYĀ : THE INDIAN JOURNAL OF STATISTICS : SERIES B

TABLE 3.4. ANALYSIS OF VARIANCE OF PERCENTAGE OF POPULATION AGED 60-ABOVE BY STRATUM WITH TWELVE VILLAGES

serial no.	stratum no.	sum of squares					mean square for error	F-ratio		
		total	party	time	party x time	error		party	time	party x time
(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	03	201.62	1.09	46.22	82.25	72.06	9.01	0.12	5.13	9.13*
2	04	523.25	94.68	0.21	16.92	411.44	51.43	1.84	0.004	0.33
3	12	268.09	15.67	85.33	0.36	166.73	20.84	0.75	0.409	0.02

*Significant at 5% level. $F5\%(1, 8) = 5.32$; $F1\%(1, 8) = 11.26$.

TABLE 4. NUMBER OF F-RATIOS SIGNIFICANT AT 5% AND 1% LEVELS CORRESPONDING TO THE FOUR DEMOGRAPHIC CHARACTERISTICS—SEX-RATIO, PERCENTAGE IN LABOUR FORCE, PERCENTAGE LITERATE, AND PERCENTAGE AGED 60-ABOVE—FOR UTTAR PRADESH (RURAL)

serial no.	number of villages in a stratum	number of strata	number of F-ratios significant at 5% level				
			P/PT	T/PT	P/E	T/E	
(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
sex-ratio							
1	4	25	1	3	-	-	-
2	8	18	0	1	0	1	2
3	12	3	1	0	0	0	0
percentage in labour force							
4	4	25	0	0	-	-	-
5	8	18	1	1*	1*	1	*
6	12	3	0	0	0	0	0
percentage literate							
7	4	25	2	0	-	-	-
8	8	18	0	1	0	1	0
9	12	3	0	0	0	0	0
percentage aged 60-above							
10	4	25	0	0	-	-	-
11	8	18	2	1	1	0	1
12	12	3	0	1	0	0	1

* Significant at 1% level.

ANALYSIS OF VARIANCE OF DEMOGRAPHIC VARIABLES

For the F -ratios party/(party \times time), in strata with 4 sample villages each, out of the total 25 for each characteristic, two for percentage literate and one for sex-ratio have come out significant at 5% level. For the 18 strata with 8 sample villages each, one F -ratio party/error for the percentage in labour force and one for percentage aged 50-above came out significant at 1% and 5% levels respectively; for strata with 12 sample villages, none of the F -ratios was significant at 5% level. The results are thus encouraging and indicate that the survey was under statistical control.*

ACKNOWLEDGEMENT

We are grateful to Dr. C. R. Rao, Head of the Department of Theoretical Research and Training, for suggestions regarding the lines of study.

*Note by the Editor

It is possible to go a step further with the data given in Table 4 by arranging them in terms of "observed" and "expected" number of F -ratios in each category. This is shown in Table 4(A) given below.

Supplementary Table 4(A) added by Editor

NUMBER OF EXPECTED AND OBSERVED F -RATIOS WHICH ARE STATISTICALLY SIGNIFICANT

serial no.	total number	number at levels of significance			
		5%		1%	
		observed	expected	observed	expected
(1)	(2)	(3)	(4)	(5)	(6)
sex-ratio					
1	50	4	2.5	0	0.50
2	90	4	4.5	0	0.90
3	15	1	0.75	0	0.15
(3.1) sub-total	155	9	7.75	0	1.55
percentage in labour force					
4	50	0	2.5	0	0.50
5	90	6	4.5	4	0.90
6	15	0	0.75	0	0.15
(6.1) sub-total	155	6	7.75	4	1.55
percentageo literate					
7	50	2	2.5	0	0.50
8	90	2	4.5	0	0.90
9	15	0	0.75	0	0.15
(9.1) sub-total	155	4	7.75	0	1.55
percentage aged 50-above					
10	50	0	2.5	0	0.50
11	90	5	4.5	0	0.90
12	15	2	0.75	0	0.15
(12.1) sub-total	155	7	7.75	0	1.55
(12.2) total	620	26	31	4	6.2

SANKHYĀ : THE INDIAN JOURNAL OF STATISTICS : SERIES B

It is clear that the observed numbers of F -ratios in each category (sex-ratio, percentage in labour force, percentage literate, percentage above 50 years) as well as for the data as a whole are less than the corresponding expected numbers at 5% and 1% levels of significance excepting for the sex-ratio at 5% level and the percentage in labour force at 1% level. This may be partly due to the fact that the F -distribution may not be the exact distribution appropriate to the ratios used here. However, the fact that the observed numbers are not widely different from but are generally only slightly less than the expected numbers makes it plausible to infer that the survey was carried out, on the whole, under statistical control.

P. C. Mahalanobis

REFERENCES

- COCHRAN, W. G. (1953): *Sampling Techniques*, John Wiley & Sons, Inc., New York.
- MAHALANOBIS, P. C. (1946): Recent experiments in statistical sampling in the Indian Statistical Institute. *J. Roy. Stat. Soc.*, 109(4), 325.
- (1958): *Lectures in Japan: Population*. Indian Statistical Institute.

Paper received: October, 1960.