International Journal of Scientific World, 5 (1) (2017) 13-17

International Journal of Scientific World

Website: www.sciencepubco.com/index.php/IJSW
doi: 10.14419/ijsw.v5i1.6611
Research paper

Factor analysis of socio-economic determinants of diseases

Shakila Khanum [1], Bilal Anjum Ansari [2], Anila Hussain [3] , Huma Qasim [4]*

[1] *Department of Gynae & Obs, Quaid-E-Azam Medical College, Bahawalpur, Pakistan*
[2] *Post Graduate Resident, Surgical Unit II, Bahawal Victoria Hospital, Bahawalpur, Pakistan*
[3] *Post Graduate Resident, Gynae & Obs. Unit I, Bahawal Victoria Hospital, Bahawalpur, Pakistan*
[4] *Statistical Assistant, Office of Executive District Officer Health, Bahawalpur, Pakistan*
**Corresponding author E-mail: huma.qasim89@yahoo.com*

Abstract

Our each and every part of life is persuaded by diversity of factors. Social and economic factors have a huge impact on person's health. The purpose of this study is to investigate the social and economic determinants of diseases by a statistical technique called factor analysis using SPSS version 21. Sample of 1500 male and female patients from public and private hospitals of Southern Punjab, Pakistan is utilized using saturation sampling technique. Results illustrate that gender, age, area living, migration of ancestors, gender of house hold head, respondent's and parent's educational level, marital status, family size and monthly income are the major social and economic determinants of diseases. Age is the major factor that affects a person's health. As age increase the chances of getting affected by disease also increase. So we can say that age and probability of getting ill are directly proportional to each other. Same is the case with income and education. Males are found to get ill more often, relatively compared to females.

Keywords: *Social Determinants; Economic Determinants; Disease; Factor Analysis; Saturation Sampling.*

1. Introduction

Our each and every part of life is persuaded by diversity of factors. Social and economic factors have a huge impact on person's health. Disease can be referred as a state in which a person's fitness becomes worse. Social factors influence standard of living, while economic factors distresses financially. Both social and economic determinants have a direct relationship with disease and person's health. Here in this manuscript Factor analysis technique has been used for determining the socio-economic causes of diseases.

Factor analysis is basically data reduction technique. Using this, data can be summarized into smaller set of variables or components. It starts with a large number of individual items/questions on a scale or measurement tool and by using analyzing techniques or factor analysis we can refine or reduce these items in to a smaller number or sub scales that measure or construct together but also measure various aspects of that construct. So we can use factor analysis to reduce a large number of related variables to more manageable number.

The impact of education, income and health behaviors on the risk of dying within the next 7.5 years with longitudinal survey study was investigated by Lantz et al in 1998. The results of cross tabulation showed that the mortality rate has a strong association with education and income. [1]

The effect of social and economic status on the individual's expected life living in Kerala, a state in India was examined by Sauvaget et al in 2011. They found that education, high income and better housing conditions contributed to long life as compared to others. [2]

A survey was conducted by Siponen et al in 2011 to study the relationship between the health of Finnish children under 12 years of age and parental socioeconomic factors (educational level, household income and working status). The analysis was done by using Pearson's Chi-Square tests, and logistic regression analysis with 95% confidence intervals. The results showed that parental socioeconomic factors were not associated with the health of children aged less than 12 years in Finland. [3]

After studying social factors affecting diseases by a questionnaire based survey in hospitals of Southern Punjab, Pakistan it was found that there was a significant relationship between social factors and diseases. They concluded that lower educational level and insufficient medical facilities in the residing area were the major factors influencing health condition. [4] There was a strong relationship between income, education and health. Health was improved if income or education increased. Stressful events and circumstances followed a socioeconomic incline, decreased as income increased. [5]

The effects of age, nativity, population size of place of residence, occupation, and household wealth on the disease and mortality experiences of Union army recruits while in service using Logistic regression were examined by Lee in 1977. The patterns of mortality among recruits were different from the pattern of mortality among civilian populations. Wealth had a significant effect only for diseases on which nutritional influence was definite. Migration spread communicable diseases and exposed newcomers to different disease environments, which increased morbidity and mortality rate. [6]

A survey in Brazil's district São Paulo, was conducted by Aranha et al in 2011 to determine the association between children's respiratory diseases reported by parents, attendance at school, parents' educational level, family income and socioeconomic status. By applying chi square test they concluded that the health of children is associated with parents' higher education, particularly mothers. Family income, analyzed according to per capita income

did not affect the number of reports of respiratory diseases from parents. [7]

Determinants can have positive or negative effect on diseases. According to Australian institute of health and welfare AIHW (2002) the factors that negatively affect a person's health are commonly known as risk factors. [8]

"The Pakistan Social and Living Standard Measurement Survey 2004-05" conducted by the Federal Bureau of Statistics was used by Arif and Naheed to determine the socioeconomic, demographic, environmental and geographical factors of diarrhea morbidity among the sampled children. Their study found a relationship between diarrhea morbidity and economic factors particularly ownership of land, livestock and housing conditions. Child's gender and age, total number of children born, mother's age and education and sources of drinking water did show significant effect on the diarrhea morbidity among children. [9]

Data from Swedish Burden of Disease study was taken by Ljung et al in 2005 to analyze the variation in health between different social and economic groups. They found that manual workers are most affected by diseases. For men cardiovascular disease and alcohol addiction, while for women depression largely contributes to health inequalities. [10]

2. Material and method

A cross sectional study was conducted at the private and public hospitals of Southern Punjab, Pakistan to collect the information from 1500 patients using saturation sampling technique. Factor analysis was done using Statistical Package for Social Science (SPSS) version 21.

Factor analysis is basically data reduction technique. Using this, data can be summarized into smaller set of variables or components. It starts with a large number of individual items/questions on a scale or measurement tool and by using analyzing techniques or factor analysis we can refine or reduce these items into a smaller number or sub scales that measure or construct together but also measure various aspects of that construct. So factor analysis is used to reduce a large number of related variables to more manageable number.

The following variables were used in the analysis:

F1. What is your gender?
F2. What is your age group?
F3. What is your current area of living?
F4. Are your ancestors' migrants?
F5. What is gender of your household head?
F6. What is your education level?
F7. What is the education level of your father?
F8. What is the education level of your mother?
F9. What is your marital status?
F10. What is your family size?
F11. What is your monthly income?
F12. Are you employed?
F13. Do you participate in social activities?
F14. Do you take bath daily?
F15. Do you smoke?
F16. Do you take exercise/play outdoor games?
F17. Do you take fruits?
F18. Do you prefer vegetables or meat?
F19. How many times a day you take meal?
F20. What is the condition of area you live in?
F21. What type of floor does your house have?
F22. What type of fuel does your household use mainly use for cooking?
F23. What is the main source of drinking water?
F24. Please indicate the disease from which you are suffering.
F25. What is your blood pressure level?
F26. Do you have medical facility in your town?
F27. Do you have medical facility in your town?
F28. How often you go to a doctor for your checkup?
F29. How many times have you been hospitalized?
F30. Is health a worry in your life?
F31. Are you suffering from stress?
F32. Do you think your home environment is suitable for your health?
F33. Do you suffer sleeping problem?
F34. Do you have feeling of hopelessness?

3. Results

Table 1

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.809
Bartlett's Test of Sphericity	Approx. Chi-Square	13869.010
	df	561
	Sig.	.000

Table 2

Communalities		
	Initial	Extraction
F1	1.000	.689
F2	1.000	.739
F3	1.000	.661
F4	1.000	.443
F5	1.000	.471
F6	1.000	.586
F7	1.000	.740
F8	1.000	.678
F9	1.000	.787
F10	1.000	.805
F11	1.000	.645
F12	1.000	.655
F13	1.000	.545
F14	1.000	.613
F15	1.000	.640
F16	1.000	.569
F17	1.000	.454
F18	1.000	.405
F19	1.000	.566
F20	1.000	.475
F21	1.000	.775
F22	1.000	.752
F23	1.000	.727

F24	1.000	.464
F25	1.000	.628
F26	1.000	.449
F27	1.000	.542
F28	1.000	.563
F29	1.000	.561
F30	1.000	.598
F31	1.000	.562
F32	1.000	.604
F33	1.000	.493
F34	1.000	.516

Extraction Method: Principal Component Analysis.

Table 3

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
F1	5.920	17.411	17.411	5.920	17.411	17.411
F2	2.430	7.148	24.559	2.430	7.148	24.559
F3	2.073	6.096	30.655	2.073	6.096	30.655
F4	1.730	5.088	35.744	1.730	5.088	35.744
F5	1.425	4.191	39.934	1.425	4.191	39.934
F6	1.302	3.831	43.765	1.302	3.831	43.765
F7	1.246	3.665	47.430	1.246	3.665	47.430
F8	1.145	3.367	50.798	1.145	3.367	50.798
F9	1.083	3.184	53.982	1.083	3.184	53.982
F10	1.036	3.047	57.029	1.036	3.047	57.029
F11	1.010	2.970	60.000	1.010	2.970	60.000
F12	.951	2.798	62.797			
F13	.947	2.785	65.582			
F14	.863	2.539	68.121			
F15	.835	2.456	70.577			
F16	.813	2.391	72.967			
F17	.783	2.302	75.269			
F18	.748	2.199	77.468			
F19	.731	2.151	79.619			
F20	.700	2.060	81.679			
F21	.695	2.044	83.723			
F22	.693	2.038	85.761			
F23	.625	1.837	87.598			
F24	.584	1.717	89.315			
F25	.550	1.617	90.932			
F26	.507	1.490	92.422			
F27	.463	1.361	93.783			
F28	.443	1.302	95.085			
F29	.403	1.185	96.270			
F30	.365	1.072	97.342			
F31	.272	.800	98.143			
F32	.246	.723	98.866			
F33	.209	.615	99.481			
F34	.176	.519	100.000			

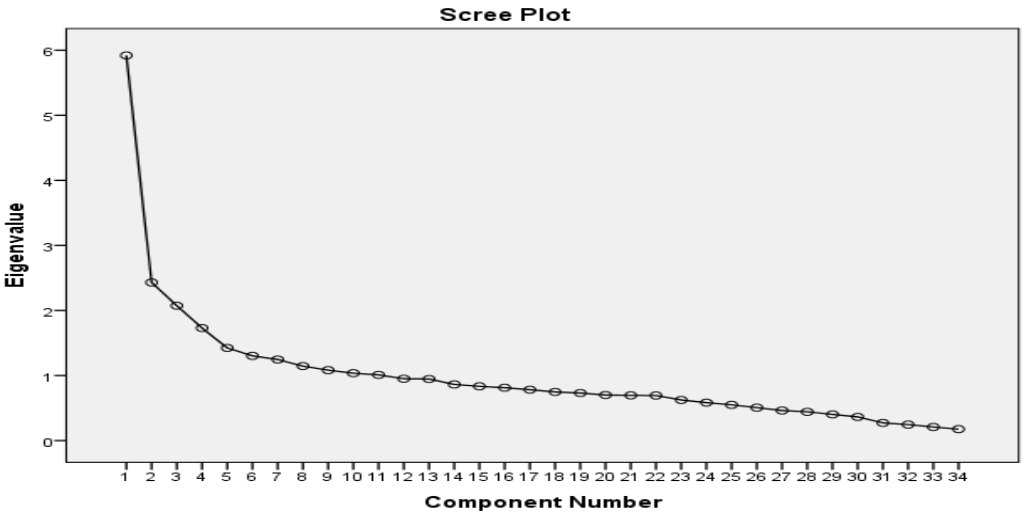

Fig. 1:

Table 4

Component Matrix[a]											
	Component										
	1	2	3	4	5	6	7	8	9	10	11
F1	.066	-.036	.686	.347	.163	-.189	.160	.046	-.010	.016	.038
F2	-.109	.816	.019	.071	-.156	.034	.114	-.022	.122	-.023	.038
F3	.715	.019	.015	-.072	-.021	-.252	.069	-.034	-.176	.136	-.157
F4	.052	-.115	-.213	.543	.196	.078	-.024	.054	.122	.127	-.089
F5	.015	-.153	.061	.359	.146	.151	.127	-.169	.013	.428	.206
F6	.662	-.186	.179	-.202	.123	.006	-.129	-.035	.043	.023	.073
F7	.701	-.162	.166	-.261	.157	.023	-.229	.034	.181	.119	.032
F8	.662	-.155	.151	-.217	.102	.029	-.241	.024	.267	.073	.013
F9	-.149	.738	.001	.033	-.194	.250	-.097	-.311	-.087	-.017	.072
F10	-.154	.041	.032	-.031	.007	-.191	.532	.478	.401	-.196	-.175
F11	.729	.200	.025	.026	-.008	.135	.049	.080	.184	.022	-.103
F12	.127	.221	.670	.160	.141	.073	-.082	-.150	-.200	-.111	-.096
F13	-.291	.075	.343	-.226	.076	-.107	.228	-.008	.112	.451	.036
F14	.571	-.138	.077	.419	.167	-.086	-.022	-.054	-.001	-.218	.000
F15	-.113	.324	.627	.287	.030	-.095	.018	.179	-.044	.001	.044
F16	.399	-.069	.218	.002	.036	.280	.010	.069	-.111	-.341	.379
F17	.536	.138	.040	-.163	.056	.207	-.041	-.084	.221	-.127	-.017
F18	-.507	-.108	.034	-.070	-.125	-.275	.010	-.186	-.049	.007	.041
F19	-.126	-.078	-.132	.336	-.046	.375	-.048	.448	.073	.230	.100
F20	.346	-.043	-.281	.376	.189	.134	.087	-.152	.139	-.064	.156
F21	.725	.165	-.137	-.012	-.271	.033	.200	.106	-.097	.005	.262
F22	.667	.167	-.111	-.023	-.321	-.006	.273	.156	-.146	-.007	.208
F23	.732	.102	.096	-.186	-.149	-.162	.174	-.012	-.159	.178	-.033
F24	-.421	-.061	.182	-.153	.083	.046	-.266	.163	.173	-.044	.295
F25	-.031	-.206	-.030	.027	.186	.454	.343	-.082	-.301	-.137	-.328
F26	.054	.464	-.232	.149	.014	-.110	-.290	.203	.074	.045	.099
F27	.593	-.008	-.242	.136	.093	-.122	.080	-.018	-.165	.155	-.179
F28	-.312	-.107	.199	-.305	-.027	.412	.269	.092	-.095	.183	.168
F29	.011	.333	-.171	-.007	.377	-.261	.230	-.264	.252	-.112	.106
F30	-.059	.311	-.193	-.178	.537	.118	.196	-.194	.089	.152	.138
F31	.080	.325	.135	-.199	.106	.343	-.101	.104	.154	-.086	-.460
F32	.136	-.001	.133	.204	-.453	.147	-.080	-.225	.297	.321	-.229
F33	.027	.339	-.104	.137	.170	-.111	-.295	.314	-.302	.122	-.121
F34	-.006	.258	-.128	-.266	.434	-.021	-.050	.248	-.279	.180	.001

Extraction Method: Principal Component Analysis.
a. 11 components extracted.

4. Discussion

The assumption for factor analysis is that KMO and Bartlett's test value must be above 0.6, and in table 1 it is 0.809, which means factor analysis is appropriate for the given data.

Another assumption is that the extracted values should be equal to or greater than 0.3, and in table 2 all the values are greater than 0.3.

Now we need to check that how many components to extract, so we need to consider a few pieces of information in the output to determine how many components met the criteria of Eigen values 1or greater than 1. We need to look at table 3.

By scanning the first column of the table 3 i.e. initial (total) only first 11 values are greater than 1, so we have to extract only those 11 factors i.e. gender, age, area living, migration of ancestors, gender of house hold head, respondent's and parent's educational level, marital status, family size and monthly income. It will keep hold on these factors and discard rest of the factors. And afterwards Eigen values will be recomputed just using those 11 factors. Basically it tells which factors extracted essentially and how much cumulative % is being explained.

We can also look at % of variance each one explained component. 1 explains 17.411 % of variance, 2 explains 7.148 % of variance. Factor 11 gives the idea of cumulative % of variance explained by this % of variance and that is 60.00. So these 11 components explain majority of the variance within this set of data.

Fig. 1 showing Scree plot gives an idea whether that was a reasonable thing to do. Now what we look from scree plot is point of inflection, so that is basically the curves. Point of inflection is being at 3 factors i.e. area living, either rural or urban. So the point of inflection is 3, which would be mean retaining or extracting 2 factors i.e. gender and age. There is also point of inflection at 13 i.e. participation in social activities, which would mean retaining factor 3.

Table 4 confirms that 11 components solution is the best option. Here 11 factors are obtained and the above whole procedure concluded that these 11 factors should be kept. So as a result we will have 11 factor's solution as shown in the table 4. These determinants of diseases are gender, age, area living, gender of house hold head, respondent's and parent's educational level, marital status, family size and monthly income that have huge influence.

Age is the major factor that affects a person's health. As age increase the chances of getting affected by disease also increase. So we can say that age and probability of getting ill are directly proportional to each other. Same is the case with income and education. If a person is well education and earns good enough, then he may avail better medical facilities as compared to those who are illiterate and earns too little income. Males are found to get ill more often, relatively compared to females. The reason behind this could be that males tend to work more than women. Usually in our society women like to stay at home and take care of their house while men works which makes his health down quickly.

References

[1] Lantz P, House J, Lepkowski J, Williams D, Mero R, Chen J, Socioeconomic Factors, Health Behaviors, and Mortality, JAMA, 279, 1998, 1703-1708. https://doi.org/10.1001/jama.279.21.1703.

[2] Saguvaget C, Socio-economic Factors & Longevity in a Cohort of Kerala State, India, Indian J Med Res, 133, 2011, 479-486.

[3] Siponen S, Ahonen R, Savolainen P, Anttila K, Children's Health And Parental Socioeconomic Factors: A Population-Based Survey In Finland, BMC Public Health, 11, 2001, 457-464. https://doi.org/10.1186/1471-2458-11-457.

[4] Shah M, Ansari A, Qasim H, Statistical Analysis of Social Causation of Diseases, International Journal of Research, 01, 2014, 980-985.

[5] Braveman P, Accumulating Knowledge on the Social Determinants of Health and Infectious Disease, Public Health Reports, 126, 2011, 126: 28-30.
[6] Lee C, Socioeconomic Background, Disease, and Mortality among Union Army Recruits: Implications for Economic and Demographic History, Explorations in Economic History, 34, 1997, 27-55. https://doi.org/10.1006/exeh.1996.0661.
[7] Aranha M, Grisi S, Escobar A, Relationship Between Respiratory Tract Diseases Declared By Parents And Socioeconomic And Cultural Factors, Rev Paul Pediatr,29, 2011, 352-356. https://doi.org/10.1590/S0103-05822011000300008.
[8] http://www.aihw.gov.au/chronic-disease-determinants/
[9] Arif A, Naheed R, Socio-Economic Determinants of Diarrhoea Morbidity in Pakistan, Academic Research International, 02, 2012, 490-518.
[10] Ljung R, SocioEconomic Differences in the Burden of Disease in Sweden, Bulletin of World Health Organization, 83, 2005, 92-99.