ADV MATH SCI JOURNAL
Advances in Mathematics: Scientific Journal **9** (2020), no.9, 6649–6658
ISSN: 1857-8365 (printed); 1857-8438 (electronic)
https://doi.org/10.37418/amsj.9.9.22 Spec. Issue on CAMM-2020

COEFFICIENT OF DETERMINATION (R^2) BY THE PATH ANALYSIS FOR THE MOST EFFECTING DEMOGRAPHIC VARIABLES ON LIFE EXPECTANCY AT BIRTH IN INDIA

VASAM NAGARAJU[1], VASANTA KUMAR V, AND K. VISWESWARA RAO

ABSTRACT. The importance of this research is being followed by R^2-value (Coefficient of determination) was evaluated through the Path analysis for the selected, most affecting demographic variables on life expectancy at birth in India. Percentage of R^2 has been calculated for variables of Infant mortality rate, Crude death rate, Crude birth rate, Total fertility rate and Under-five mortality rate. The primary data has been collected through the sources of United Nations Population Division (UNPD), from the year 1960 to 2018. In this research Path analysis techniques utilized and Path coefficients (Standardized coefficients) found with the help of Multiple Regression analysis. By Applying Step-wise regression procedure built the structural models; the result is significant at 1 percent los. The percentage of R^2 value is highest in Net effects by CBR (28.02); Joint or interaction effect was found (X_3X_4) CDR and CBR with (49.34). Conclude that the life expectancy at birth in India or region or any other country has been influenced by most of the demographic factors. Utilized SPSS software for the Analysis.

1. INTRODUCTION

Expectation of life is defined as the estimated average number of years, a person is expected to live beyond certain age when he is subjected to the current risk

[1]*corresponding author*

2010 *Mathematics Subject Classification.* 62N02, 97K80.

Key words and phrases. Life expectancy at birth, Coefficient of determination, Path analysis and India.

of mortality at different ages. Goran (2020), investigated [1] the effect of the socioeconomic development on life expectancy at birth as an indicator of mortality or longevity in five EU accession candidate countries (Macedonia, Serbia, Bosnia and Herzegovina, Montenegro, and Albania). Using aggregate time series pool data on an annual level from UN and World Bank databases for the period 1990-2017 and Full Information Maximum Likelihood model, it was found that this connection between the socioeconomic conditions and life expectancy at birth is a prerequisite for longer life in all these five countries. Nurmawati and Kismiantini (2018), investigated [2] the relationships among age, sex, weight, height, smoking behaviour, and blood pressure on health status of adults in Indonesia. They constructed the path analysis using the secondary data of the fifth wave of the Indonesian Family Life Survey in 2014/2015. Their hypothesized model suggested that age, sex, weight, height, and smoking behaviour had an effect on blood pressure and that all variables influenced health status. Abihishek Singh et al. (2017) used [4] the data from the Sample Registration System (SRS) for the period 1981-2011 and examined the trends in inequality in length of life in India and 15 major states of India by decomposing the inequality in length of life into the contributions of age and causes of death. They observed that while gains in life expectancy have been remarkably steady both overall and across states of India, gains against life span variance have scarcer.

John Cantiello et al. (2015) aimed [5] to study why young adults between the ages of 18 and 24 are the largest uninsured population subgroup. Their investigation shown that coverage by a private health insurance plan in the 2005 sample was largely a matter of having a higher socioeconomic status and having a non-minority status. In 2008 each of the attitudinal variables [6] (socioeconomic status, minority status, perceived health, perceived value, and perceived need) emerged as the significant predictors of private insurance.

Ewards and Shripad Tuljapurkar (2005), observed [7] that the second half of the twentieth century witnessed substantial convergence in life expectancy around the world. Jean Marie Robine and Karen Ritchie (1991) reviewed [8] and evaluated the usefulness of healthy life expectancy as a global indicator of changes in a population's health. They have taken the data for some western countries for the decade 1981-1990 and computed the healthy life index. Erriss (1988) explained, [9] social indicators are used to monitor the social system, it helps to identify changes and to guide intervention to alter the course of social

change. Life expectancy at birth is also key indicator for the development of social system and human life style. Hugh Leobner and Edwin D.Driver (1973) studied [10] the relation of several demographic and socio-economic variables to fertility and use of contraception through multivariate analysis techniques.

1.1. **MATERIALS.** The data available through primary sources was utilized. Data on life expectancy at birth, Infant mortality rate, Under-five mortality rate, Crude birth rate, Crude death rate and Total fertility rate in India year wise, has been taken from the United Nations Population Division (UNPD).

2. METHODS

Statistical and Mathematical methods utilized in the study are Measure of averages and variation, useful tests of inference, Student's t-test and Anova procedure, correlation and multiples correlation coefficient, multiple Regression (step-wise) procedure and model building for the data set (linear) and Path analysis techniques. Computer software, SPSS utilized for generation of findings with these statistical methods.

2.1. **PATH ANALYSIS.** The method of path analysis builds on ordinary multiple regression analysis. It applies only to sets of relationship among the variables which are linear, additive and causal. In multiple regression each predictor variable has a direct effect on the response variable. However, variables may also affect the response variable through one or more intervening variables. Path analysis is a technique for analysing such causal relationships. Nagaraju et al. (2018) examined, [3] the technique of path analysis which was developed during 1920s by Sewall Wright as an aid to the quantitative development of genetics gained popularity in social science studies with the further expositions made by Duncan and Land.

3. RESULTS AND DISCUSSION

Using the $L.E_(T)$ at birth, IMR, CDR, CBR, TFR and U5MR for the time period, 1960-2018, the demographic variables are presented in the following table-1. Life expectancy at birth (Total), Infant mortality rate, Crude death rate, Crude birth rate, Total fertility rate and Under-five mortality rate in India.

Table 1: The values of the demographic variables from 1960-2018

Year	LE_T	IMR	CDR	CBR	TFR	U5MR
1960	41.42	161.4	22.18	42.00	5.91	242.1
1961	42.03	159.1	21.65	41.75	5.90	238.5
1962	42.64	156.9	21.13	41.48	5.89	235.3
1963	43.25	154.9	20.61	41.20	5.88	232.3
1964	43.87	153.1	20.10	40.90	5.86	229.5
1965	44.50	151.3	19.59	40.59	5.83	226.8
1966	45.14	149.6	19.09	40.27	5.79	224.2
1967	45.78	148.0	18.60	39.97	5.75	221.7
1968	46.43	146.4	18.12	39.67	5.70	219.1
1969	47.08	144.7	17.65	39.39	5.65	216.4
1970	47.74	142.9	17.19	39.11	5.59	213.6
1971	48.40	141.1	16.73	38.82	5.52	210.6
1972	49.06	138.9	16.28	38.52	5.44	207.2
1973	49.72	136.6	15.84	38.20	5.36	203.5
1974	50.37	134.0	15.41	37.87	5.28	199.4
1975	51.01	131.2	14.99	37.53	5.19	194.9
1976	51.63	128.1	14.60	37.21	5.11	189.9
1977	52.22	124.8	14.23	36.91	5.03	184.6
1978	52.79	121.3	13.89	36.64	4.96	179.2
1979	53.32	118.0	13.58	36.40	4.89	173.6
1980	53.81	114.7	13.29	36.17	4.83	168.2
1981	54.27	111.6	13.03	35.92	4.77	163.1
1982	54.69	108.7	12.79	35.63	4.70	158.3
1983	55.07	105.9	12.56	35.29	4.64	153.9
1984	55.44	103.4	12.34	34.89	4.56	149.7
1985	55.80	100.9	12.12	34.42	4.48	145.7
1986	56.17	98.4	11.89	33.89	4.40	141.8
1987	56.55	95.9	11.65	33.32	4.31	137.8
1988	56.96	93.4	11.40	32.73	4.22	133.8
1989	57.40	90.9	11.14	32.12	4.13	129.9
1990	57.87	88.6	10.86	31.52	4.05	126.2

Table 1 continued from previous page

Year	LE_T	IMR	CDR	CBR	TFR	U5MR
1991	58.35	86.3	10.59	30.92	3.96	122.6
1992	58.85	84.2	10.32	30.35	3.88	119.3
1993	59.35	82.1	10.06	29.79	3.80	116.0
1994	59.84	80.1	9.81	29.26	3.72	112.8
1995	60.32	78.0	9.58	28.75	3.65	109.4
1996	60.78	75.8	9.37	28.26	3.58	106.0
1997	61.23	73.5	9.18	27.79	3.51	102.4
1998	61.67	71.2	9.00	27.32	3.45	98.8
1999	62.09	68.9	8.84	26.86	3.38	95.2
2000	62.51	66.6	8.69	26.40	3.31	91.6
2001	62.91	64.3	8.56	25.94	3.24	88.0
2002	63.30	62.1	8.43	25.49	3.18	84.5
2003	63.70	59.9	8.31	25.03	3.11	81.1
2004	64.10	57.8	8.19	24.57	3.04	77.7
2005	64.50	55.7	8.07	24.09	2.97	74.4
2006	64.92	53.6	7.96	23.56	2.90	71.1
2007	65.35	51.5	7.84	23.00	2.82	67.9
2008	65.79	49.4	7.72	22.39	2.74	64.6
2009	66.24	47.3	7.60	21.76	2.66	61.4
2010	66.69	45.1	7.49	21.11	2.58	58.2
2011	67.13	43.0	7.39	20.50	2.51	55.1
2012	67.55	40.9	7.31	19.92	2.44	52.1
2013	67.93	38.9	7.25	19.42	2.38	49.1
2014	68.29	36.9	7.21	18.98	2.33	46.3
2015	68.61	35.0	7.19	18.63	2.30	43.6
2016	68.90	33.2	7.20	18.33	2.27	41.1
2017	69.17	31.5	7.21	18.08	2.24	38.7
2018	69.42	29.9	7.23	17.86	2.22	36.6

It is observed from table-1, that the life expectancy at birth has shown significant increase. The IMR, CDR, CBR, TFR and U5MR are decreasing significantly. Where $Y_1 = L.E_T, X_2 = IMR, X_3 = CDR, X_4 = CBR, X_5 = TFR$ and $X_6 = U5MR$.

Using SPSS programmed software for the data of table-1, the multiple correlation coefficients between the demographic variables life expectancy at birth (Total), Infant mortality rate, Crude birth rate, Crude death rate, Total fertility rate and Under-five mortality rates of India, 1960-2018, was obtained and presented in Table 2.

TABLE 2. Correlation Matrix

Variable	LE_T	IMR	CDR	CBR	TFR	U5MR
LE_T	1.000	-0.996	-0.985	-0.979	-0.991	-0.996
IMR	-0.996	1.000	0.971	0.988	0.998	1.000
CDR	-0.985	0.971	1.000	0.932	0.959	0.972
CBR	-0.979	0.988	0.932	1.000	0.995	0.987
TFR	-0.991	0.998	0.959	0.995	1.000	0.998
U5MR	-0.996	1.000	0.972	0.987	0.998	1.000

The Table 2, represents the association between the life expectancy at birth and the selected demographic variables is observed to have negative direction for all. The highest correlation was given by IMR and U5MR with -0.996 and -0.996 respectively. That means the life expectancy at birth was most affected by the infants and under-five mortality rates in India.

Using SPSS software from the data of Table 1, the following multiple regression equations fitted. The structural equations have been developed with the help of Step-wise Regression procedure.

Model-I, is given as,

$$LE_T = 75.879 + (IMR) * (-0.201) + e.$$

Model-II, is given as,

$$LE_T = 77.119 + (IMR) * (-0.138) + (CDR) * (-0.585) + e.$$

Model-III, is given as,

$$\begin{aligned} LE_T = {} & 83.539 + (IMR) * (-0.026) + (CDR) * (-0.918) \\ & + (CBR) * (-0.414) + e. \end{aligned}$$

Model-IV, is given as,

$$\begin{aligned} LE_T = {} & 81.951 + (IMR) * (-0.083) + (CDR) * (-0.908) \\ & + (CBR) * (-0.568) + (TFR) * (2.802) + e. \end{aligned}$$

The Unstandardized and Standardized Coefficients; t-statistic and significant values of the Regression models have been presented in Table 3.

TABLE 3. Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t-statistic	Sig-value
		B	Std. Error	Beta		
1	(Constant)	75.879	0.258		294.620	0.000
	IMR	-0.201	0.003	-0.996	-79.875	0.000
2	(Constant)	77.119	0.191		402.957	0.000
	IMR	-0.138	0.006	-0.686	-22.579	0.000
	CDR	-0.585	0.056	-0.319	-10.510	0.000
3	(Constant)	83.539	0.866		96.506	0.000
	IMR	-0.026	0.015	-0.131	-1.707	0.093
	CDR	-0.918	0.059	-0.501	-15.462	0.000
	CBR	-0.414	0.055	-0.382	-7.509	0.000
4	(Constant)	81.951	0.988		82.930	0.000
	IMR	-0.083	0.025	-0.412	-3.363	0.001
	CDR	-0.908	0.056	-0.495	-16.209	0.000
	CBR	-0.568	0.075	-0.524	-7.573	0.000
	TFR	2.802	0.987	0.417	2.840	0.006

The statistical model adopted for this is a simple linear model as given: Considered for the four variables, Regression model - 4.

$$Y_1 = P_{12}X_2 + P_{13}X_3 + P_{14}X_4 + P_{15}X_5$$

$R^2_{1.2345}$ = Coefficient of determination =0.999; (from the model summary).

The R^2, values shown 0.999 is highest for the both Regression model-3 and model-4. Which represent best fit models.

Standardized coefficients from table-3 are considered as path coefficients. (P12 = -0.412, P13 = -0.495, P14 = -0.524 and P15 = 0.417)

R_u = Residual Variable = $\sqrt{(1 - R21.234)} = 0.0316$

$R^2 = P^2_{12} + P^2_{13} + P^2_{14} + P^2_{15} + 2P_{12}P_{13r23} + 2P_{12}P_{14r24}$
$+2P_{12P15r25} + 2P_{13P14r34} + 2P_{13P15r35} + 2P_{14P15r45}$.

The component analysis of coefficient of determination $R^2_{1.2345}$. Net effects due to X_2, X_3, X_4 and X_5 and Joint effects,R^2 and Percentage of R^2 values have been calculated in Table 4. The R^2 value has been taken from model summary; Anova table results shown significant for the fitted models.

TABLE 4. Component Analysis of Coefficient of determination, $R^2_{1.2345}$

Details	Component	Amount	R^2 (percentage)	Percentage of R^2
Net effect				
Due to X2	P212	0.169744	16.9744	17.3208163
Due to X3	P213	0.245025	24.5025	25.0025510
Due to X4	P214	0.274576	27.4576	28.0179591
Due to X5	P215	0.173889	17.3889	17.7437755
Total of Net Effects		0.863234	86.3234	88.0851020
Joint effect				
Due to X2X3	2P12P13r23	0.39605148	39.6051	40.4133673
Due to X2X4	2P12P14r24	0.42659468	42.6594	43.5300000
Due to X2X5	2P12P15r25	-0.34292078	-34.2920	-34.9918367
Due to X3X4	2P13P14r34	0.48348432	48.3484	49.3351020
Due to X3X5	2P13P15r35	-0.39590397	-39.5903	-40.3982653
Due to X4X5	2P14P15r45	-0.45142752	-45.1427	-46.0639795
Total of Joint Effects		0.11587821	11.5879	11.8243877
Total multiple Determination		0.97911221	98	100

From Table 4, the net effects with respective to X_2, X_3, X_4 and X_5 are observed to be 16.97, 24.50, 27.45 and 17.38 respectively. The highest percentage of R^2 was given by the variables X_3 (CDR) and X_4 (CBR). Total of net effects percentage of R^2 is 88.08. Interaction effects have been calculated for $X_2X_3, X_2X_4, X_2X_5, X_3X_4, X_3X_5$ and X_4X_5 with the help of components of $2P_{12}P_{13r23}, 2P_{12P14r24}, 2P_{12P15r25}, 2_{P13P15r35}$ and $2P_{14P15r25}$. The corresponding percentage of R^2, 40.41, 43.53, -34.99, 49.33, -40.39 and -46.06 respectively. Total

of joint effects is given by 11.82 and the highest percentage of R^2 is 49.33 with (X_3X_4).

4. Conclusion

The fitted multiple regression models are shown significant. The increase of life expectancy at birth in India or state or any region is dependent on the decrease of Infant mortality rate, Under-five mortality rate, Crude birth rate, Crude death rate and Total fertility rates. The highest percentage of Coefficient of Determination (R^2) was given by CBR (28.02) among all the Net effects. In the Interaction effects, the best set of factors found in CDR and CBR (49.34) for the set X_3X_4 . Suggested, the well-being of a country or a region, demographic variables are the key indicators.

References

[1] G. Miladinov: *Socio-economic development and life expectancy relationship: evidence from the EU accession candidate countries*, Genus Journal, **76**, art.no. 2, (2020). https://doi.org/10.1186/s41118-019-0071-0

[2] W.P. Nurmawati, Kismiantini: *Path analysis for determining health factors in Indonesia*, Journal of Physics: Conf: Series, **1320** (2019), 012018. doi: 10.1088/1742-6596/1320/1/012018

[3] N. Vasam, V. Kumar, V. Rao: *Joint and Net effect on life expectancy at birth through the literacy rate and Infant mortality rate of India and state-wise by Path analysis*, ARPN Journal of Engineering and Applied Sciences, **13**(7) (2018), 2588-2593.

[4] A. Singh, A. Shukla, F. Ram, K. Kumar: *Trends in equality in length of life in India: a decomposition analysis by age and causes of death*, Genus, **73**, art.no 5, 2017. doi: 10.1186/s41118-017-0022-6

[5] J. Cantiello, M.D. Fottler, D. Oetjen, N. Jackie Zhang: *The impact of demographic and perceptual variables on a young adult's decision to be covered by private health insurance*, BMC Health Services Research **15**, art.no. 195, 2015. doi: 10.1186/s12913-015-0848-6

[6] R.D. Edwards, S. Tuljapurkar: *Inequality in life spans and a new perspective on mortality convergence across industrialized countries*, Population and development review, **31**(4) (2005), 645-674.

[7] R. Veenhoven: Happy life-expectancy, Social indicators Research, **39** ((1996), 1-58. https://doi.org/10.1007/BF00300831

[8] J. Marie Robine, K. Ritchie: *Healthy Life expectancy: evaluation of global indicator of change in population health*, British Medical Journal 1991;302:457. doi: https://doi.org/10.1136/bmj.302.6774.457

[9] A.L. FERRISS: *The uses of social indicators, Oxford University Press*, Social forces, **66**(3) (1988), 601-617. DOI: 10.2307/2579568

[10] H. LEOBNER, E.D. DRIVER: *Differential Fertility in Central India: A Path Analysis*, Demography, **10** (1973, 329-350. doi: 10.2307/2060843

DEPARTMENT OF MATHEMATICS
KONERU LAKSHMAIAH EDUCATION FOUNDATION
ANDHRA PRADESH, INDIA.
E-mail address: nagaraju.stats@gmail.com

DEPARTMENT OF MATHEMATICS,
KONERU LAKSHMAIAH EDUCATION FOUNDATION, VADDESWARAM, GUNTUR (DT)
ANDHRA PRADESH - 522 502, INDIA.
E-mail address: vvkumar@kluniversity.in

DEPARTMENT OF BIO-STATISTOCS
(RETD.) NIN,
HYDERABAD, INDIA.
E-mail address: visweswararao.stats@gmail.com