IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 12, Issue 1 Ver. I (Jan. - Feb. 2016), PP 66-70
www.iosrjournals.org

Relationship Between Average Personal Income And Genetic Factors Of Patients In Abia State University Teaching Hospital, Aba, Nigeria.

[1]George, Uchechukwu &[2] Iwuagwu, Ben Chioma
Department of Statistics, Michael Okpara University of Agriculture, Umudike, Abia State, Nigeria

Abstract: *This study investigated the relationship between average personal income and genetic factors in Abia State University Teaching Hospital, Aba. The data for the study was a secondary data which constitute a sample of 500 patients. The data was observed to be normally distributed. The basic assumptions underlying the classical linear regression model which includes multi-co-linearity, autocorrelation and hetroscedasticity was verified and a linear regression model of average personal income and genetic factors was fitted. In conclusion, it was recommended for the general public to be enlightened on their average income management and the adverse health effect emanating from genetic factors so as to maintain a healthy living.*
Keywords: *Average personal income, Genetic factors, Multiple Regression Model and Hetroscedasticity*

I. Introduction

People in higher income groups tend to have better physical health. This relationship is well established throughout history, across geographical boundaries, and for almost every disease and condition (Adler et al., 1994; Antonovsky, 1967). Though it may seem obvious that severe poverty might erode physical health through the effects of poor nutrition, crowded and dirty living conditions, and inadequate medical care, the association exists across the income range. Not only do those just above the poverty line have better health than those in poverty, but those in the highest income levels have better health than those just below them (Marmot et al., 1991). In addition, the relationship cannot be explained by lack of access to health care, and it exists even in populations with universal access to medical care (Adler and Snibbe, 2003). The effect is greatest for the poorest groups, yet in the United States it has the greatest impact on the middle classes because the largest numbers of people are in those income ranges (Adler and Snibbe, 2003). Adler et al. (1994) summarized findings to data about the relationship between income and health, and examined possible explanations for the basis of the association. The first possibility they considered was that the relationship results from common underlying genetically based factors. They suggested that, for example, physical size or intellectual capacity could contribute directly to both income and physical health, resulting in a spurious association between the two effects. They dismissed this possibility as unlikely, however, pointing out that the association between job status and health persists after adjustment for height and body mass index (Singer et al., 1999), and that intellectual capacity does not appear to be reliably linked to health. In addition, they noted that any genetic predispositions involved would probably be important only when environmental and behavioral factors impinged on them. Since then, most research on possible mechanisms accounting for the relationship has focused on another explanation suggested by Adler et al. (1994): that income influences biological functions that in turn affect health status. Generally, the process that subsequent research has begun to articulate (Adler and Snibbe, 2003) centers around the increasing demands and decreasing resources for dealing with those demands associated with lower levels of in- come. This results in greater exposure to stress at lower income levels, as well as in greater psychological response to that stress.

II. Literature Review

Stephen et al (1998) investigated the relationship between body mass index and blood pressure among university students in Maiduguri, Nigeria. In their work a sample of convenience was used to recruit participant for the study data on gender height, weight, body mass index and blood pressure were obtained using a research development data from the variables under study which was measured using a standardized procedure. A total of 351 students participated out of which 248 were males (70.7%) and 103 were females (29.3%) with mean weight obtained respectively the mean SBP and DBP were computed. Their study revealed significant correlation between BMI and BP for males and no significant correlation for females.

Nadia et al (2009) carried out an investigation on decreasing association between body mass index and Blood Pressure over time. In their study two independent cross-sectional examination survey conducted in 1989 and 2004 from age 25-64 years. They found out a linear relationship between blood pressure and body mass index was marked in 2004.

Masazumi et al, (1996) investigated that adolescent's blood pressure is associated with body weight and body mass index. They review the relationship between interviewing changes in blood pressure. In their study they revealed a correlation between systolic blood pressure and body mass index, the non-linear least squares procedure was used to estimate weight function with a sample of 420 participant. Their study use the double exponential gompertz model obtaining the first and second derivative of the resulting prediction to compute acceleration and velocity of weight.

Abdul and Marwan (2013) made a research on the effect of interest rate, inflation rate, and GDP on real economic growth in Jordan over the period 2000-2010. Regression analysis was conducted to test growth rate with interest rate which showed that current interest rate has an influence power on growth rate. Also, regression was used to test growth rate with inflation rate; it showed that inflation rate has influence power on economic growth rate. Finally regression used to test GDP, interest rate, and inflation rate together; according to them, results have shown that current GDP and one lag GDP have influence power to economic growth rate.

III. Methodology

The data used in this research work is secondary and primary data collected from Abia State University Teaching Hospital, (ABSUTH) Aba, on genetic factors of patients admitted in the hospital for 500 patients. The genetic factors considered here are; Age, Height, weight, body temperature, genotype, blood group and blood pressure. And questionnaire distributed on average personal income.

Concept of regression

Regression Analysis is concern with the study of dependence of one variable, the dependent variable on one or more other variables,called the explanatory variables with the aim of estimating and predicting the population or average value of the former in terms of the known or fixed values of the later.

Assumptions of Classical Linear Regression Model (CLRM)

1. The regression model must be linear in parameters but may not be linear in variables.
2. The error or disturbance term e_i follows a normal distribution with mean zero and constant variance i.e $e_i \sim N(0, \sigma^2)$
3. The independent sample are fixed in repeated sampling i.e. cov $(x_i e_i)$ =0 the explanatory variable and error term are independent.
4. There is no autocorrelation between the error term i.e.

Cov$(e_i e_j / x_i x_j)$ = Cov$(e_i e_j)$ = 0 if x is non stochastic.

5. There is no collinearity between the explanatory variable that is the

explanatory variables has no relationship.

General Regression Model

The estimated regression model is obtained by minimizing the error sum we have,

$$\widehat{Y_1} = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} \ldots + \hat{\beta}_K X_{Ki} \qquad (1)$$

Putting the above normal equation in matrix form we have

$$\begin{pmatrix} \sum Y_i \\ \sum X_{2i} Y_i \\ \vdots \\ \sum X_{ki} Y_i \end{pmatrix} = \begin{pmatrix} n & \sum X_{2i} & \cdots & \sum X_{ki} \\ \sum X_{2i} & \left(\sum X_{2i}\right)^2 & \cdots & \sum X_{2i} \sum X_{ki} \\ \vdots & \vdots & \vdots & \vdots \\ \sum X_{ki} & \sum X_{ki} \sum X_{2i} & \cdots & \left(\sum X_{ki}\right)^2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} \qquad (2)$$

Which is $X^T Y = (X^T X)(\beta)$

To solve for β, we make it the subject of the formula in equation (1.2) by multiplying both side by the inverse of $(X^T X)$ i.e

$(X^T X)^{-1} (X^T Y) = \beta (X^T X) (X^T X)^{-1}$

hence $\beta = (X^T X)^{-1} (X^T Y)$ (3)

Where $(X^T X)^{-1} = \frac{1}{\det(X^T X)} adj(X^T X)$

where

$$\mathbf{X^TY} = \begin{pmatrix} \sum Y_i \\ \sum X_{2i} Y_i \\ \vdots \\ \sum X_{ki} Y_i \end{pmatrix}, \quad (\mathbf{X^TX}) = \begin{pmatrix} n & \sum X_{2i} & \cdots & \sum X_{ki} \\ \sum X_{2i} & \left(\sum X_{2i}\right)^2 & \cdots & \sum X_{2i} \sum X_{ki} \\ \vdots & \vdots & \vdots & \vdots \\ \sum X_{ki} & \sum X_{ki} \sum X_{2i} & \cdots & \left(\sum X_{ki}\right)^2 \end{pmatrix}$$ and

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{pmatrix} \qquad (4)$$

IV. Result and Discussions

This chapter discusses the results of the research.

Regression Equation

AV. IN = 169429 + 127.273 AGE + 266.764 WEIGHT - 84.2087 HEIGHT - 1312.85 TEMP. + 184749 D1 - 235.022 D2 - 2572.14 D3 (5)

Coefficients of regression

Term	Coef	SE Coef	T	P	95% CI	VIF
Constant	169429	211513	0.8010	0.423	(-246151, 585009)	
AGE	127	579	0.2199	0.826	(-1010, 1264	1.05152
WEIGHT	267	270	0.9887	0.323	(-263, 797)	1.05318
HEIGHT	-84	516	-0.1631	0.871	(-1099, 931)	1.00883
TEMP.	-1313	5736	-0.2289	0.819	(-12583, 9958)	1.01142
D1	184749	11376	16.2397	0.000	(162397, 207102)	1.02179
D2	-235	5278	-0.0445	0.965	(-10605, 10135)	1.00575
D3	-2572	5921	-0.4344	0.664	(-14206, 9062)	1.00916

Summary of Model

S = 74136.1 R-Sq = 35.59% R-Sq(adj) = 34.67%

PRESS = 2.788179E+12 R-Sq(pred) = 33.58%

Shows the regression analysis of average personal incomes on age, weight, height, temperature, coded blood pressure(D1),coded blood groups(D2) and for coded genotypes(D3). The result of the table revealed that the coefficients of age, weight and blood pressure have positive signs. The implication of this is that both weight and blood pressure are positively related to average personal income. The table also showed that the coefficients of height, temperature and genotype and blood group have negative signs which means that they are negatively related to average personal income. The table equally revealed that of all the variable coefficients in the model only the coefficient of weight and blood pressure are statistically significant. Table 4.1 also showed that the R-square 0.03559 that is about 35.59 percent of the total variation observed in the dependent variables was accounted for by the independent variables included in the model.

Analysis of Variance

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	7	1.49396E+12	1.49396E+12	2.13424E+11	38.831	0.000000
AGE	1	7.89196E+08	2.65750E+08	2.65750E+08	0.048	0.826048
WEIGHT	1	1.18689E+10	5.37233E+09	5.37233E+09	0.977	0.323310
HEIGHT	1	4.42001E+09	1.46121E+08	1.46121E+08	0.027	0.870544
TEMP.	1	8.42043E+09	2.87903E+08	2.87903E+08	0.052	0.819063
D1	1	1.46741E+12	1.44949E+12	1.44949E+12	263.727	0.000000
D2	1	1.98513E+07	1.08973E+07	1.08973E+07	0.002	0.964502
D3	1	1.03712E+09	1.03712E+09	1.03712E+09	0.189	0.664192
Error	492	2.70411E+12	2.70411E+12	5.49616E+09		
Lack-of-Fit	484	2.70276E+12	2.70276E+12	5.58421E+09	33.092	0.000008
Pure Error	8	1.35000E+09	1.35000E+09	1.68750E+08		
Total	499	4.19807E+12				

Durbin-Watson Statistic

Durbin-Watson statistic = 2.04029

The d-value obtained indicates that there is no evident of autocorrelation in the residuals of the observation

Hetroscedasticity Test

In this section, the breusch pagan godfrey (BPG) test would be used to test for constant variance in the residuals of the observations in line with the stated procedures in previous section of the chapter above.

$\widehat{\sigma^2} = \frac{\sum \widehat{u_i^2}}{n} = \frac{2.70411\,E+09}{500} = 5408.220.$

And $P_i = \frac{u_i^2}{\widehat{\sigma^2}}$

Then regress the P_i on the uncorrelated explanatory variables

Then the regression equation of

$P_i = \propto_1 + \propto_2 Z_{2i} + \ldots + \propto_m Z_{mi}$ is given as

Regression Equation

pi = -9.65088e+006 + 36299.7 AGE + 9078.71 WEIGHT - 10367.5 HEIGHT + 251087 TEMP. + 128017 D1 + 352248 D2 - 355001 D3 (6)

Analysis of Variance of the BPG test

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	7	8.96577E+13	8.96577E+13	1.28082E+13	0.537	0.806877
AGE	1	2.52872E+13	2.16176E+13	2.16176E+13	0.906	0.341739
WEIGHT	1	7.09417E+12	6.22237E+12	6.22237E+12	0.261	0.609883
HEIGHT	1	1.94737E+12	2.21485E+12	2.21485E+12	0.093	0.760790
TEMP.	1	1.06374E+13	1.05310E+13	1.05310E+13	0.441	0.506863
D1	1	1.97446E+12	6.95956E+11	6.95956E+11	0.029	0.864489
D2	1	2.29612E+13	2.44792E+13	2.44792E+13	1.026	0.311708
D3	1	1.97560E+13	1.97560E+13	1.97560E+13	0.828	0.363397
Error	492	1.17439E+16	1.17439E+16	2.38698E+13		
Lack-of-Fit	484	1.17432E+16	1.17432E+16	2.42628E+13	261.695	0.000000
Pure Error	8	7.41712E+11	7.41712E+11	9.27140E+10		
Total	499	1.18336E+16				

Following the Anova table, of the above regression equation, the explained sum of square is given as ESS = 8.96577e13 Hence Θ = $\frac{1}{2}(ESS)$ =4.48288e13.The tabulated chi-square value is given as $\chi^2_{m-1} = \chi^2_{1-\propto,m-1}$,Where $\propto = 0.05$ and m=7 (the number of independent variables) therefore, $X^2_{0.95\,,}$ 6= 15.508.

Conclusion: Since Θ-value exceeds chi-square value, i.e 4.48288e13 > 15.508. We reject H_0 and conclude that there is heteroscedasticty in the residuals of the observation.

Coefficients of the BPG model

Term	Coef	SE Coef	T	P	95% CI	VIF
Constant	-9650880	13938998	-0.69237	0.489	(-37038185, 17736426)	
AGE	36300	38144	0.95166	0.342	(-38645, 111244)	1.05152
WEIGHT	9079	17782	0.51057	0.610	(-25859, 44016)	1.05318
HEIGHT	-10367	34035	-0.30461	0.761	(-77239, 56504)	1.00883
TEMP.	251087	378020	0.66422	0.507	(-491646, 993821)	1.01142
D1	128017	749721	0.17075	0.864	(-1345033, 1601067	1.02179
D2	352248	347836	1.01269	0.312	(-331179, 1035676)	1.00575
D3	-355001	390215	-0.90976	0.363	(-1121694, 411693)	1.00916

Summary of Model

S = 4885671 R-Sq = 0.76% R-Sq(adj) = -0.65%

PRESS = 1.192078E+16 R-Sq(pred) = -0.74%

Durbin-Watson Statistic

Durbin-Watson statistic = 2.01637

V. Findings

- The data for the analysis satisfy the assumption of multicolinearity by measure of the VIF was applied thereby reducing the effect of removing or addition of variables
- The Durbin-Watson's test indicates that there is no autocorrelation in the data positive or negative
- The heteroscedasticity was tested by Breausch Pagan Godfrey(BPG) test and was found present in the data but was corrected by the method of Weighted Least Square (WLS).
- The overall significant test indicates that there is a significant relationship between average personal income and weight and also with blood pressure respectively.
- The coefficient of determination (R^2) indicates that the regression model is adequate since it approaches to one.

VI. Summary

The regression analysis result shows that AVI is the dependent variable while age, weight, height, temperature, blood pressure, blood group and genotype are the explanatory variables, looking at the regression equation below.

$$\text{AV.IN} = 169429 + 127.273\ \text{AGE} + 266.764\ \text{WEIGHT} - 84.2087\ \text{HEIGHT} - 1312.85\ \text{TEMP.} + 184749\ \text{BP} - 235.022\ \text{B.GROUP} - 2572.14\ \text{GENOTYPE} \quad (4.3)$$

From the model we observe that when the explanatory variables are all zeros the constant term is feasible. Also a unit change in average personal income will lead to 127.273 change in age in positive direction also a unit change in average personal income will lead to 266.764 change in height positively while a unit change in average personal income will lead to 84.2087 change in height negatively, 1312.85 change in temperature in a negative manner, 184749 change in blood pressure positively and 235.022 and 2572.14 change in blood group and genotype in a negative manner respectively.

VII. Conclusion

From the result of the previous chapter, conclusion can be drawn from this study since there exist a positive relationship among blood pressure, age and weight this implies that a unit change in age would lead to a unit change in blood pressure also a unit change in weight would also result to a unit change in the average personal income therefore there is a high tendency of being wealthy as one grows older and also being hypertensive with a maximum increase in weight to become overweight.

References

[1]. Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., Folkman, S., Kahn, R. L., and Syme, S. L. (1994). Socioeconomic status and health. Am. Psychol. Pp.49:15-24.
[2]. Singer, B. H., and Ryff, C. D. (1999). Hierarchies of life histories and associated risks. In N. E. Adler, M. Marmot, B. S. McEwen and J. Stewart (eds.), Socioeconomic Status and Health in Industrial Nations. New York: New York Academy of Sciences, pp. 96-115.
[3]. Adler, N. E., and Snibbe, A. C. (2003). The role of psychosocial processes in explaining The gradient between socioeconomic status and health. Curr. Directions Psychol. Sci. 12(4), 119-123.
[4]. 4. Stevens, J.,Cai,J., Pamuk, E.R.,Willamson,D.F.,Thun, M.J and wood,J.L.(1998). The effect of age on the association between body mass index and mortality.
[5]. New England Journal of Medicine. Pp.338. (1):1-7
[6]. Naidia, A.N., Neela, J. and Rai, N.P. (2009), maternal body mass index and birth weight.
[7]. Nutrition News, National Institute of Nutrition, Hyderabad. Pp12 (2):12-132
[8]. Masazium .A, Midori .S, Randolph L. C. (1996), Correlation between systemic Blood pressure, and physical development in adolescence. An American Journal of epidemiology John Hopkins University.
[9]. Abdul and Marwan (2013).The effect of Interest Rate, Inflation Rate, GDP, On Real Economic Growth rate in Jordan. Asian Economic and Financial Review, 2013, 3(3):341-354.
[10]. Minitab statistical software(2013) version 17.