

Multiple Logistic
Regression Modelling Of
Drug Abuse/ Exposure
Among The People Of
Kirinyaga Central, Kenya

Gideon Lang'at

Pure and Applied Sciences, Kirinyaga University

Josephine Njeri Ngure

Pure and Applied Sciences, Kirinyaga University

Kinyanjui Jeremiah

Pure and Applied Sciences, Kirinyaga University

Gitonga Harun

Pure and Applied Sciences, Kirinyaga University

ABSTRACT

The research study is to determine the relationship between drug abuse and the major factors that lead to drug abuse putting into consideration gender, and to ascertain if they really lead to drug abuse in the area. These factors include; lack of awareness on the effects of drugs, poverty, unemployment, early child exposure to drugs and other toxic substances, failure of the government agencies, peer pressure, lack of child support centers and rehabilitation centers. The study finds how drug abuse affects social life and economic development of the people of Kirinyaga central through which statistical induction would translate to other sub-counties in Kirinyaga County, the objective of the study is to find the relationship between drug abuse and poverty and unemployment among people of Kirinyaga Central. The research employ a survey design targeting all wards in Kirinyaga central where data was collected from the residents. Stratified sampling technique was used to select two areas for the study from each strata/ward. The sampled areas were stratified according to the size of the population and simple random sampling was used to select 357 respondents from which the data was obtain through questionnaires. Prior to data collection, a pilot study was conducted to ascertain the reliability and validity of the instrument.

Keywords: drug abuse, exposure, NACADA, logistic regression, Akaike Information Criterion, Null deviance

1. INTRODUCTION

Drug is any substance which when introduced into the body will alter the normal biological and psychological functioning of the body, (Escandon and Galvez, 2006), and drug abuse is the chronic use of a drug for a reason other than for which it is intended, that is bad use of a drug for example alcohol, Mandrax, cocaine etc. (IJSTR, 2013). Exposure is the condition of being exposed or uncovered from the effects of the drugs that impair with the health of an individual.

Frustrations and emotional stress due to failures, sorrows of miseries of life lead people to drug abuse, and from Robert Merton's theory of anti-social behavior emphasizes that an individual finds no avenue towards achievement of his goals he may be compelled to indulge in anti-social behavior and one way of doing that is by retreating from social participation through drug addiction.

Annual of internal medicine from American college of physician's files states that lack of basic needs and economic factors such as unemployment are the common factors leading many people to taking drugs.

Drug abuse has become a major torn to social life and economic development among the people of Kirinyaga central and research done by NACADA and World Bank (2012) shows that the number of people affected by the drugs are increasing day by day despite the effort by the government to gap the increase. This has raised big concern and the government through NACADA has tried to put more effort to ensure that the problem is curbed.

The situation demand for a rehabilitation center and strengthening guidance and counseling sessions in schools for the students.

2. REVIEW OF PREVIOUS STUDIES

Center for Suicide Research and Prevention and the Department of Social work and Social Administration, Hong Kong (2011) done a research on Drug Abuse Among youths ad Family Relationships and they found that the drug prevalence for in-school adolescent is about 5% to 7% for boys and 3% to 4% for girls also parents' divorce, separation or passing away has high influence on substance abuse among adolescents and they finally conclude that influence from family on substance abuse (early age exposure) play a major role to drug abuse, a more related research was also done by kwong Leung at. el(2006) a research on Drug Abuse situation among Ethnic Minorities in Hong Kong he found that there is high average frequency of drug abuse (19.2 times per week currently and 23.9 times per six months before) they indicate that even though drug abusers realized the health risk, they remain addicted to drugs.

National Survey on Drug use and Health, Columbia and united states (2013) find out that in 2013 despite the fact that those who are unemployed are the ones who indulge more in drugs 65.8% who were employed abuse drugs and 53.8 of those who are unemployed take drugs,

Masudi, Omar B (2011) also did a research on factors influencing drug abuse among the youth in Mombasa County, Kenya find out that peer pressure, curiosity, availability of drugs and cost also plays a major role during inception for drug abuse.

3. METHODOLOGY

3.1 DATA COLLECTION

The research was on an unbiased and valid sample collected from four Kirinyaga county wards (Kanyekini, Inoi, Mutira, kerugoya) in Kenya which was to be a representative of the whole population. The target group for collection of data was a family head or representative who could avail the required information as was needed. Questionnaires were administered to appropriate respondents who could provide necessary response to the set questions. There was also use of interviews in case of aged and semi-illiterate respondents and call backs in case of missing information. No secondary data was used in this research.

3.2 REVIEW OF MULTIPLE LOGISTIC REGRESSION MODEL

Multiple logistic regression is a regression model in which the response variable takes only two possible binary values 0 or 1, we use logistic regression equation to predict the probability of dependent variable taking the dichotomy / binary / categorical value 1 or 0 or more generally, the presence or absence of an attribute of interest and normally there are two or more measurement variables.

[John Wiley et al (1996)] multiple logistic regression equation is given as;

$$\hat{p} = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)} \quad (1)$$

Where \hat{p} is the expected probability that the outcome is present and $x_i, i = 1, 2, \dots, k$ are the independent variables β_0 through β_k are the regression coefficients.

Sometimes, the multiple logistic regression model is written differently as

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = (\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k) \quad (2)$$

James H. Stepleton (2009) shows that logistic regression model is heteroskedastic and the response variable is not normally distributed this is because the error variance differs for each value of the predicted score

Logistic regression is a special case of generalized linear model and thus analogous to linear regression and has the following assumptions.

- The condition $y|x$ is a Bernoulli distribution rather than Gaussian distribution, because the dependent variable is binary.
- The predicted values are the probabilities and they are restricted to (0, 1) through the logistic distribution function because logistic regression predicts the probability of a particular outcome (Fishers 1936).

Odds ratio – is a measure of association between a certain factor in a population. It tells you how the presence or absence of one factor has an effect on the presence of another factor (e.g. how poverty has made people to engage in drug abuse).

Considering the case of drug abuse and poverty,

Odds of success on drug abuse is

$$p/(1 - p) , p \text{ is the probability of success}$$

Odds of success on poverty is

$$r/(1 - r) , r \text{ is the probability of success}$$

Therefore, the odd ratios

$$OR = \frac{\frac{p}{1-p}}{\frac{r}{1-r}} \quad (3)$$

Deviance - deviance is basically a measure of how much unexplained variation there is in our logistic regression model, the higher the value the less accurate the model

$$D_{Null} = -2 \ln\left(\frac{\text{likelihood of null model}}{\text{likelihood of a saturated model}}\right)$$

$$D_{fitted} = -2 \ln \left(\frac{\text{likelihood of fitted model}}{\text{likelihood of a saturated model}} \right) \tag{4}$$

Likelihood Ratio Test - Likelihood ratio test is used to access the model fit and one way of doing this is to compare deviance of the predictor model with that of the null model on a chi-square distribution with degrees of freedom equal the number of predictor variables.

4. DATA ANALYSIS AND RESULTS

The first thing was to check whether the data is follows a normal distribution which was done by fitting a normal q-q plot to the data. The fitted regression line was as shown in the figure meaning the data followed the normal distribution as shown in figure 1.

The collected data was analyzed using R statistical software. From the contingency table on poverty it was observed that a larger proportion of people who responded say YES to poverty i.e. $248/334 = 0.74251497$ from the results it implies that about 74.251497% of people indulge in drug abuse because of poverty, and from the correlation test it was observed that poverty is 0.6784243 this implies that poverty is much correlated to drug abuse also the p-value for poverty is $1.167e-07$ this is approximately zero implying that drug abuse is dependent on poverty.

The study established that unemployment is one of the major factor that lead many people in Kirinyaga central to drug abuse, this implies that a larger percentage of the people who are unemployed are likely to indulge in drug abuse compared to those who are employed. From the correlation test it was found that unemployment is strongly correlated to drug abuse with value 0.6813623 and the p-value is $9.794e-08$ this implies that drugs abuse is dependent on unemployment level. There was also found that there was high positive correlation between drug abuse / exposure and poverty with a correlation value of 0.7. The correlation tests shows that unemployment, and poverty are strongly correlated to the response variable drug abuse with 68%. And their 95% confidence interval is much narrower (between 0.49-0.81) this shows that the data were not dispersed.

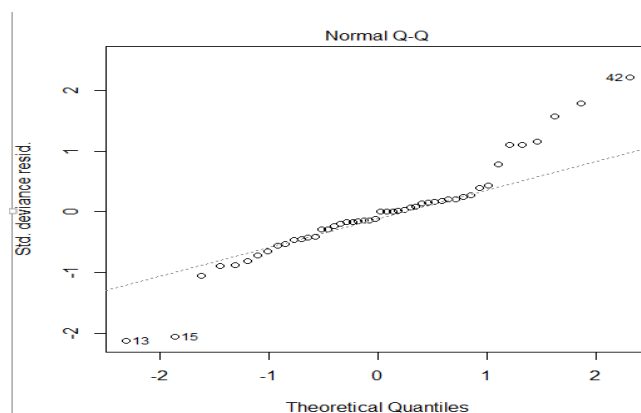


Figure 1

The above graph is not bell-shaped this shows that the errors are normally distributed in the data and there is normality in the assumptions (i.e. data was normally distributed and come from the same distribution/sample).

From the contingency table on poverty it was observed that a larger proportion of people who responded said YES to poverty i.e. $248/334 = 0.74251497$. From the results it implies that about 74.251497% of people indulge in drug abuse because of poverty. The probability of NO to drugs is $82/286$ and the probability of saying YES to drugs because of unemployment is $204/286=0.713286713$ this implies that there is high probability of someone taking drugs because of unemployment.

Likelihood ratio test - was used to access the model fit by comparing deviance of the predictor model with that of the null model, If the predictor model has significantly smaller deviance (chi-square using difference in degrees of freedom of the two model) then we conclude that there is significant association between the 'predictor' and 'outcome' Neyman et al (1933) and Michael J. Crawley(2015), Model selection to obtain the best fit of the model that best shows the relationship between the factors and the response variable (drug abuse) was done

$$LR = -2(L_0 - L_1) \quad (5)$$

Wald Statistics – was used to test whether the variable is making a significant contribution to the prediction of outcome and it was found that poverty and unemployment were statistically significantly related to drug abuse / exposure in Kirinyaga County.

$$W = \frac{\beta}{se(\beta)} \sim N(0,1) \quad (6)$$

The model that was selected was the one that had the lowest AIC which showed that it was the best fit for the data obtain from the study after a series of repetitive analysis.

The fitted model was as shown

$$\ln\left(\frac{\hat{p}}{(1-\hat{p})}\right) = -6.1551 + 0.4085(unemployment) + 0.3467(poverty) \quad (7)$$

The log odds of taking drugs is 0.4085 times higher in persons who are unemployed than those who are employed. Taking the antilog of the regression coefficients, exponential $(0.4085) = 1.5$, which is actually the crude or unadjusted odds ratio. This means that the odds of taking drugs are 1.5 times higher among unemployed people compared to employed persons. Also, the odds of taking drugs are 1.4 times higher among poor people compared to those living above poverty level. The odds ratio was greater than 1 which showed that poverty is associated with higher odds of drug abuse. Szumilas (2010).

4. CONCLUSION AND RECOMMENDATION

I recommend more research and application to be done using other statistical method and also in other wards of the county so as to come up with a better representation of drug abuse / exposure in Kirinyaga County as a whole. There should be more job creations and this would help people of Kirinyaga central to abandon the use of drugs and concentrate on their work. NACADA should invest more on training so that people would be enlightened on different drugs and their related effect. The government should put more effort to eradicate illicit brews in the village. In school's guidance and counselling session should be strengthened so that students and teenagers are enlighten early enough to prevent them from indulgence.

5. REFERENCES

- Szumilas M. Explaining Odds Ratios. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*. 2010; 19 (3):227-229.
- Research Report on A Study on the Drug Abuse Situation among Ethnic Minorities in Hong Kong Submitted to Research Subcommittee of the Action Committee Against Narcotics By Kwong-leung Tang Hung Wong Chau-kiu Cheung Department of Social Work The Chinese University of Hong Kong In Collaboration with Unison Hong Kong—For Ethnic Equality June 2006.
- Neyman, Jerzy; Pearson, Egon S. (1933). "On the Problem of the Most Efficient Tests of Statistical Hypotheses" (PDF). *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. **231** (694–706): 289–337.
- Long J. Scott (1997). *Regression Models for Categorical and Limited Dependent variables* (second Edition). 14:35-98 Thousand Oaks, CA: sage publications
- Linear Statistical Models, 2nd Edition James H. Stapleton ISBN: 978-0-470-23146-3 474 pages September 2009
- INTRODUCTION TO STATISTICAL MODELLING IN R P.M.E.Altham, Statistical Laboratory, University of Cambridge. January 7, 2015
- Michael J. Crawley(2015). An introduction using R 2, (2):53-211,273-279
- Sampling Techniques, 3Rd Edition Paperback – November 21, 2007by William G. Cochran
- Cox, D. R.; Hinkley, D. V. (1974). *Theoretical Statistics*. Chapman and Hall. ISBN 0-412-12420-3.
- An Introduction to Categorical Data Analysis: 1st (First) Edition Hardcover – February 28, 1996 by John Wiley, Vic Barnett (Editor), Ralph A. Bradley (Editor) Alan Agresti
- Allen M, Donohoe WA, Griffin A, Ryan D, Mitchell-Turner MM. comparing the influence of parents and peers on the choice to use drugs. *Criminal Justice abd Behaviour*. 2003; 30:163-186