Statistical Analysis of Sampled Cases of Sexually Assaulted Female Students in a Nigerian Private University

Olusegun Adelodun[1,*], Tejumade Adekola[2], Damilola Adekanmbi[3]

[1]Institute of Education, Obafemi Awolowo University, Ile-Ife, Nigeria
[2]Bowen University, Iwo, Nigeria
[3]Department of Statistics, Ladoke Akintola University, Ogbomoso, Nigeria

Abstract The study was carried out on sexually assaulted females and some relationships such as parents, lecturers, open places, among others and also males guilty of sexual assault and some reasons such as alcohol, pornography, among others in Bowen University, Iwo, Nigeria. This aimed at determining the relationship between sexual assault and factors responsible for it, specifically to formulate statistical model for the problem and determining significant factors responsible for sexual assault. The study adopted regression analysis with the assumption that errors are uncorrelated among others. Logistic regression was employed through the help of SPSS Version 15.0. The sample size of 125 derived from stratified sampling making use of six faculties in Bowen University as the strata. Neyman allocation formula was applied. Result obtained gave a logistic regression model for the data set. The study concluded that the number perpetrated by their relatives and number of assaults in an open place had contributed immensely to the rave of sexual assault among females in a Nigerian University, especially in Bowen University.

Keywords Sexual Assault, Logistic Regression Analysis, Neyman Allocation, Female Students, Bowen University

1. Introduction

In sampling cases, the employed tool has been the statistical analysis. The reason is not far-fetched; statistical analysis is the science of collecting, exploring and presenting large amounts of data to discover underlying patterns and trends. Statistics are applied every day – in research, industry and government – to become more scientific about decisions that need to be made. Logistic regression analysis being one of the most widely used statistical techniques employed in almost every field of application.

Logistic regression is one of the varieties of popular multivariate tools used in biomedical informatics. It is one of the most common models for prediction and has been applied to sexual assault [1, 2]. From previous studies, logistic regression is widely used in medical literature especially for correlating the dichotomous outcomes with the predictor variables that include different physiological data. In logistic regression, the predicted odd ratio of positive outcome is expressed as a sum of product. Product is formed by multiplying the values of independent variable and its coefficients. The probability of positive outcome is obtained from the odd ratio through a simple transformation. The problems are formulated first from the logistic regression. Then, the coefficient obtained from the logistic regression is used to calculate the predictor variables [3].

Logistic regression is used for prediction by fitting data to the logistic curve. It requires the fitted model to be compatible with the data. In logistic regression, the variables are binary or multinomial. Multinomial Logistic Regression analysis is capable of showing the best way to find conclusion and be made as parsimonious model to describe the relationship between dependent and independent variables. Binary Logistic Regression is one of the logistic regression analysis methods whereby the independent variables are dummy variables. Independent variables consist of different size levels whereas dependent variables must be linear and fulfills the response that is needed for this method. A logistic regression model is the result of non-linear transformation of the linear regression model. The difference between logistic regression and linear regression is that the outcome variable in logistic regression is dichotomous [4].

This study uses data to determine the relationship between sexual assault and factors that are responsible for it. The study addresses the following two objectives: (i) to formulate statistical model for the problem; and (ii) to predict the

factors responsible for sexual assault.

The United Nation defines violence against women as any act of sexual assault that results in or is likely to result in physical, sexual, or mental harm or suffering to women including threats of such acts, coercion, or arbitrary deprivation of liberty whether occurring in public or private life [5, 6]. Following the adoption of the definition above, different agencies of United Nation at different conferences had resolved and recommended many actions to the nations on addressing various aspects of violence (including sexual and physical violence) and human rights violations against women [7-9].

The World Health Organization (WHO) multi-country study on women health and domestic violence against women clearly demonstrates that, violence against women is wide-spread and deeply ingrained, and has serious impacts on women's health and well being and its continued existence is morally indefensible [10, 11]. Using operational definitions of types of violence against women in multi-country World Health Organization Study, violence against women when committed by an intimate partner such as husband or/and other family members is referred to as domestic violence, while that perpetrated by people outside this group is called nonpartner violence [10, 12, 13]. Both types of violence affect the physical, psychological, and social status of the female gender and therefore constitute threat to the health of women [10, 12, 13].

The situation of violence against women in Nigeria may not be different from reports from other countries [14-16]. However, there is dearth of National data on the subject. The available local reports indicate that violence against women, both partner and nonpartner, are common [14, 17]. Sexual violence against women gained more relevance because of its capability to promote the spread of HIV/AIDS [18-20].

This study statistically reviewed the sampled cases of sexually assaulted female students in a Nigerian private university. The findings may further bring the issue to public attention with the overall goal of reducing the problem of sexual assault against students in other campuses.

2. Materials and Methods

2.1. General Information about the University

Bowen University, Iwo, Osun State, South-western Nigeria, is a private institution owned by the Nigerian Baptist Convention and named after Reverend Thomas Jefferson Bowen who pioneered Baptist work in Nigeria in 1850. The University obtained a final Government approval in 2001 but commenced academic activities on the 4th November 2002 through its first student intake of about 500.

The University is conceived as a centre of learning and research of distinction; combining academic excellence with love of humanity, borne out of a God-fearing attitude, in accordance with the Baptist tradition of ethical behaviour, social responsibility and democratic ethos. The guiding philosophy of the University is "Excellence and Godliness".

This emphasizes the importance the University attaches to the attainment of excellence imbued with Godliness in all its programmes and training. This motto is the principle underpinning all academic pursuits in the University.

The focal point of the logo is the open book. The sun stands for light, life and energy. The palm frond, a dominant foliage of the Nigerian rainforest belt, used for celebrations in the Christian tradition, symbolizes freshness and a new beginning. The wheel embodies the unending and unbroken circle of knowledge informed by Godliness, with the Baptist steering humanity along the path of knowledge.

The vision of the University is to be one of the foremost institutions for producing first rate graduates in science, technology and the humanities, who combine godliness and academic excellence with a strong sense of ethical responsibility, towards the development of the society at large.

The mission is to provide high quality, morally sound and socially relevant education, in the most cost-effective manner, to all its students, irrespective of gender, race, colour, ethnicity or religion. The University will develop, promote and be run as a unique Christian institution with the following distinctive features:

i) Unquestionable Christian morality;
ii) Vocational and apprenticeship orientation towards the full employment of its products; and
iii) A work-study programme of activities to instill work ethics an self-reliance in students.

The University currently runs a College and five Faculties as follows:

College of Health Sciences, Faculty of Agriculture, Faculty of Science and Science Education, Faculty of Social and Management Sciences, Faculty of Humanities, and Faculty of Law. These students will serve as our subject for the study.

2.2. Sampling Method

Stratified Sampling adopting Neyman allocation was used for the study. The purpose of the method is to maximize survey precision, given a fixed sample size. With Neyman allocation, the 'best' sample size for stratum i would be:

$$n_i = n \left(\frac{N_i \sigma_i}{\sum_{i=1}^{N} N_i \sigma_i} \right) \tag{1}$$

where n_i is the sample size in stratum i,
   n is total sample size,
   N_i is the population size in stratum i, and
   σ_i is the standard deviation in stratum i.

2.3. Data

The sample used for this study was 125 students. 9 from faculty of Agriculture, 35 students were from Science and Science Education, 50 from Social and Management Sciences, 12 from Humanities, 12 from Medical Sciences and 7 from the faculty of Law. Although, 150 questionnaires

were administered to both males and females but 125 were returned.

The data was collected from January to April 2014 by the researcher. The researcher administered the questionnaire on 125 randomly selected students. The following thirteen variables are considered for classifying the presence of sexually assaulted (SA).

The independent variables are number perpetrated by your friend (H1), number perpetrated by your partner (H2), number perpetrated by your parent (H3), number perpetrated by your relative (H4), number perpetrated by your pastor / imam (H5), number perpetrated by your lecturer (H6), number perpetrated by your stranger (H7), number of assaults in a secluded place (H8), number of assaults in an open place (H9), taking alcohol (M1), watching pornography (M2), joining peer groups (M3) and weak governmental laws (M4).

Table 1.   Description of Variables

Variable	Definition	Characteristic
SA	Have you been sexually assaulted?	No = 0, Yes = 1
H1	Number perpetrated by your friend	Ordinal
H2	Number perpetrated by your partner	Ordinal
H3	Number perpetrated by your parent	Ordinal
H4	Number perpetrated by your relative	Ordinal
H5	Number perpetrated by your pastor / imam	Ordinal
H6	Number perpetrated by your lecturer	Ordinal
H7	Number perpetrated by a stranger	Ordinal
H8	Number of assaults in a secluded place	Ordinal
H9	Number of assaults in an open place	Ordinal
M1	Taking alcohol	Ordinal
M2	Watching pornography	Ordinal
M3	Joining peer groups	Ordinal
M4	Weak governmental laws	Ordinal

The data description of the thirteen independent variables and the dependent variable are provided in Table 1. Data analysis is performed using SPSS (2007), V15.0, SPSS Inc.

2.4. Logistic Regression Analysis

Logistic regression analysis is a popular and widely used analysis that is similar to linear regression analysis except that the outcome is dichotomous (e.g., yes/no). In essence, we examine the odds of an outcome occurring (or not), and by using the natural log of the odds of the outcome as the dependent variable the relationships can be linearized and treated much like multiple linear regression.

Multiple logistic regression analysis applies here, since there is a single dichotomous outcome and more than one independent variable. The outcome in logistic regression analysis is often coded as 0 or 1, where 1 indicates that the outcome of interest is present, and 0 indicates that the outcome of interest is absent. If we define p as the probability that the outcome is 1, the multiple logistic regression model can be written as follows:

$$\hat{p} = \frac{\exp(b_0 + b_1 X_1 + b_2 X_2 + \ldots + b_p X_p)}{1 + \exp(b_0 + b_1 X_1 + b_2 X_2 + \ldots + b_p X_p)} \qquad (2)$$

$\hat{p}$ is the expected probability that the outcome is present; X_1 through X_p are distinct independent variables; and b_0 through b_p are the regression coefficients. The multiple logistic regression model is sometimes written differently. In the following form, the outcome is the expected log of the odds that the outcome is present, $In\left(\frac{\hat{p}}{1-\hat{p}}\right)$

$$\text{and} \quad In\left(\frac{\hat{p}}{1-\hat{p}}\right) = b_0 + b_1 X_1 + b_2 X_2 + \ldots + b_p X_p \quad (3)$$

Notice that the right hand side of the equation above looks like the multiple linear regression equation. However, the technique for estimating the regression coefficients in a logistic regression model is different from that used to estimate the regression coefficients in a multiple linear regression model. In logistic regression the coefficients derived from the model (e.g. b_1) indicate the change in the expected log odds relative to a one unit change in X_1, holding all other predictors constant. Therefore, the antilog of an estimated regression coefficient, $\exp(b_1)$, produces an odds ratio.

2.5. Logistic Regression Model

Several tests had been performed in the logistic regression analysis. The tests are model fitting test, parameter estimation and classification. Model fitting test is to check whether all the variables are suitable to be used in the logistic regression. Model fitting test is done by using likelihood ratio statistic. Likelihood ratio is defined as

$$LR[i] = -2\big(LL(a) - LL(a, B)\big) \qquad (4)$$

where $LL(a)$ is the log-likelihood of the beginning model and $LL(a, B)$ is the log-likelihood of the ending model. Likelihood ratio is distributed chi-square with i degree of freedom.

Parameter estimation is to estimate each independent variable that contributes to the sexual assault. Parameter estimation is done using log-odd ratio. The log-odd ratio is defined as

$$B = In\left[\frac{P(Y=m)}{P(Y=0)}\right] = \beta_0 + \sum_{p=1}^{n} \beta_p \, x_p \qquad (5)$$

where $m = 1, 2, \ldots, n$.

To find the log-odd ratio, the probability of each event is calculated. Odds ratio measure the incidence when the independent variable increases by one unit. The odds ratio is defined as

$$\frac{P(Y=m)}{P(Y=0)} = \exp\big(\beta_0 + \sum_{p=1}^{n} \beta_p \, x_p\big) \qquad (6)$$

Classification is to predict the respondents in sexual assault. From the calculated coefficients, the probability of each sample is calculated. The probability is defined as

$$P(Y = m) = \frac{\exp(g(x))}{1 + \sum \exp(g(x))} \qquad (7)$$

while for the reference category;

$$P(Y = 0) = \frac{1}{1 + \sum \exp(g(x))} \qquad (8)$$

3. Results and Discussion

The Data analyses will be performed on sexually assaulted people based on the logistic regression procedure as follows:

Table 2 shows the case processing summary. There are 125 cases used in the analysis.

Table 2. Case Processing Summary

Unweighted Cases		N	Percent
Selected Cases	Included in Analysis	125	100.0
	Missing Cases	0	0
	Total	125	100.0
Unselected Cases		0	0
Total		125	100.0

From Table 3, given the base rates of the two decision options (80/125 = 64% decided to continue the research, 36% decided to allow it to stop), and no other information, the best strategy is to predict, for every case, that the subject will decide to continue the research. Using that strategy, you would be correct 64% of the time.

Table 3. Classification Table

			Predicted		
			Decision		Percentage
Observed			Stop	Continue	Correct
Step 0	Decision	Stop	0	45	0
		Continue	0	80	100.0
	Overall Percentage				100.0

Under Table 4 we see that the intercept-only model is ln(odds) = .575. If we exponentiate both sides of this expression we find that our predicted odds [Exp(B)] = 1.778. That is, the predicted odds of deciding to continue the research is 1.778. Since 80 of our subjects decided to continue the research and 45 decided to stop the research, our observed odds are 80/45 = 1.778.

Table 4. Variables in the Equation

		B	S.E.	Wald	Df	Sig.	Exp(B)
Step 0	Constant	.575	.186	9.534	1	.002	1.778

Under Model Summary in Table 5, we see that the -2 Log Likelihood statistic is 105.756. This statistic measures how poorly the model predicts the decisions -- the smaller the statistic the better the model. The Cox & Snell R^2 can be interpreted like R^2 in a multiple regression, but cannot

reach a maximum value of 1. The Nagelkerke R^2 can reach a maximum of 1.

Table 5. Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	105.756	.369	.506

The variables in the equation output in Table 6 shows us that the regression equation is

$$In(ODDS)$$
$$= -1.703 - .108H1 - .430H2 + .236H3$$
$$+ 1.111H4 - .763H5 - .074H6 + .172H7$$
$$+ .306H8 + .546H9 - .045M1 + .500M2$$
$$- .134M3 - .547M4.$$

Table 6. Variables in the Equation

Term	B	S.E.	Wald	Df	P-value	Exp(B)
Constant	-1.703	1.443	1.393	1	.238	.182
H1	-.108	.218	.247	1	.619	.897
H2	-.430	.225	3.651	1	.056	.651
H3	.236	.288	.670	1	.413	1.266
H4	1.111	.446	6.198	1	.013	3.038
H5	-.763	.488	2.444	1	.118	.466
H6	-.074	.208	.126	1	.723	.929
H7	.172	.166	1.071	1	.301	1.188
H8	.306	.215	2.033	1	.154	1.358
H9	.546	.222	6.037	1	.014	1.727
M1	-.045	.430	.011	1	.916	.966
M2	.500	.536	.871	1	.351	1.649
M3	-.134	.291	.213	1	.644	.875
M4	-.547	.411	1.771	1	.183	.182

Furthermore, the significant values for H4 and H9 are lower than 0.05. The results in Table 6 show the p-values for each variable. The variables that had significant values below 0.05 are considered to be taken into the logistic regression model. The variables that had values below than 0.05 are H4 and H9. Since there are multiple predictors (such as, H1, H2, H3, H5, H6, H7, H8, M1, M2, M3 and M4) without a statistically significant association with the response, the reduced model is therefore given as:

$$In(ODDS) = -1.703 + 1.111H4 + .546H9.$$

Since we do have categorical variables in our design (for male and female), we can use this model to predict the odds, hence we have the following steps.

We can now use this model to predict the odds that a subject of a given gender will decide to continue the research. The odds prediction equation is

$$ODDS = e^{a + b_1 X_1 + b_2 X_2}.$$

If our subject is a female (gender = 0), then the

$$ODDS = e^{-1.703 + 1.111(0) + .546(0)} = e^{-1.703} = 0.182.$$

That is, a female is only 0.182 as likely to decide to continue the research as she is to decide to stop the research. If our subject is a male (gender = 1), then the

$$ODDS = e^{-1.703 + 1.111(1) + .546(1)} = e^{-0.046} = -0.044.$$

We can easily convert odds to probabilities. For our female,

$$\hat{Y} = \frac{ODDS}{1 + ODDS} = \frac{0.182}{1.182} = 0.154.$$

That is, our model predicts that 15% of female will decide to continue the research. For our men, we can perform the same computation similarly.

It is pertinent to state that the vast majority of cases of sexual assault against female university students in Nigeria go unreported for various reasons associated with victim shaming, stigma, character assassination, public backlash and limited access to justice for victims. In many cases, female students who have reported such cases have been subsequently targeted for reprisal attacks by thugs, cultists or university teachers.

The mass abuse of female university students in Nigeria is fueled by the lack of a consistent and clear policy by university governing bodies and school authorities concerning sexual assault of female students. Student handbooks and codes of conduct for staff and students are generally 'silent' on this and do not adequately address issues of violence against women. Sexual assaults are hardly mentioned during orientation for first year students neither are there any dedicated channels of redress or support for students who experience this form of violence whilst on campus.

It is not enough for university authorities to continue to disassociate themselves, single out, disown or suspend individual perpetrators of such grievous crime of gender based violence and abuse of public office and trust without providing any sort of support and care for the victims and their families.

Sexual assault is a crime under the criminal and penal code in Nigeria. Sexual assault is a clear violation of Article 3 (4) of the protocol to the African Charter on Human and Peoples rights on the Rights of Women in Africa (2003), Article 2 (d) of The Convention on Elimination of all forms of Discrimination Against Women (CEDAW) 1979 and Section 24 of the Violence Against Person Prohibition (VAPP) Act 2014. Sexual assault of female students directly impedes on the progresses made in advancing Girl-Child education, in fulfillment of the Millennium Development Goals (MDGs) and newly adopted Sustainable Development Goals (SDGs), as well at the objectives sited on the AU Agenda 2063 regarding Girl-child education.

4. Conclusions

No one knows better than surgeons how multiple factors can combine to produce patient outcomes. Logistic regression analysis is a powerful tool for assessing the relative importance of factors that determine outcome. A basic understanding of logistic regression analysis is the first step to appreciating both the usefulness and the limitations of the technique. The problem of sexual assault against female student has become a social problem. It is a one which affects all classes of female students in our society. Female students of all colours, races, economic classes, languages and religions are not spared. The available research findings show that it is a persistent problem which needs persistent confrontation. This study justifies that relationship between victim and perpetrator such as friends, partner, parents, lecturer, stranger, pastor/imam, secluded place, alcohol, pornography, sexual peers and weak governmental laws on the sexually assaulted females leading to the conclusion that the number perpetrated by their relatives and number of assaults in an open place have contributed significantly to the act of sexually assaulting female students in Nigerian Universities, in particularly Bowen University, Iwo.

REFERENCES

[1] A.M. Young, M. Grey and C.J. Boyd. Adolescents' experiences of sexual assault by peers: prevalence and nature of victimization occurring within and outside of school. J Youth Adolesc. 2009 Sep; 38(8): 1072-83.

[2] H. Cecil and S.C. Matson. Sexual victimization among African American adolescent females: examination of the reliability and validity of the Sexual Experiences Survey. *J Interpers Violence.* 2006 Jan; 21(1): 89-104.

[3] M. Zhou. Revisiting Ethnic Entrepreneurship: Convergencies, Controversies, and Conceptual Advancements. *International Migration Review*, 2004.

[4] D. W. Hosmer and S. Lemeshow. Applied Logistic Regression 2000, English, Article edition

[5] Declaration on the elimination of violence against women. New York, NY: United Nations. United Nation General Assembly resolution document. RES/48/104; 1993. p. 1-10. [Last accessed on 2010 Oct 20].

[6] R. Jewkes, C. Garcia-Moren and P. Sen. Sexual violence. In: World report on violence and health. Geneva: World Health Organization; 2002. p. 149-81.

[7] J.C. Caldwell. The International Conference on Population and development, Cairo, 1994. Is its plan of action important, desirable and feasible? The nature of the forum. Forum. Health Transit Rev 1996; 6: 71-122.

[8] United Nations. Fourth World Conference on Women, Beijing. Platform of Action on Violence against Women. Beijing, China. United Nations publication; 1995. p. 1-6.

[9] United Nations General Assembly. United Nations Millennium Declaration. Millennium Development Goals (MDG). New York. United Nations publication; 2000. p. 1-9.

[10] C. Garcia-Moreno, H.A. Jansen, M. Ellsberg, L. Heise and C. Watts. WHO multi-country study on women's health and domestic violence against women: Initial results on

prevalence, health outcomes and women's responses. Geneva: World Health Organization; 2005. p. 1-50.

[11] C. Garcia-Moreno and C. Watts. Violence against women: An urgent public health priority. Bull World Health Organ 2011; 89: 2.

[12] A.L. Ilika, P.I. Okonkwo and P. Adogu. Intimate partner violence among women of childbearing age in a primary health care centre in Nigeria. Afr J Reprod Health 2002; 6: 53-5.

[13] N. Amah, S.O. Shittu and M.A. Abdul. Risk scoring for domestic violence in pregnancy. Niger J Clin Pract 2008; 11: 18-21.

[14] Amnesty International. Nigeria Rape the silent weapon. Abuja. Amnesty International publication. 2006. p. 3-42.

[15] O.I. Fawole, A.J. Ajuwon and K.O. Osungbade. Evaluation of interventions to prevent gender-based violence among young female apprentices in Ibadan, Nigeria. Health Education, 2005 105 (3): 186-203.

[16] F.I. Omorodion and O. Olusanya. The Social Context of reported rape in Benin City Nigeria. Afr J Reprod Health 1998: 11; 37-43.

[17] C.O. Aimakhu, O. Olayemi, C.A. Iwe, F.A. Oluyemi, I.E. Ojoko, K.A. Shoretire, *et al*. Current causes and management of violence against women in Nigeria. J Obstet Gynaecol 2004: 24; 58-63.

[18] C. Ugbuji. Violence against Women Impact on Their Reproductive Health. Trop Obstet Gynaecol 2004; 21: 61-64.

[19] H.O. Caroline and A. Richter. Exploring Intersections between Gender Violence; Lessons from Zimbabwe. Afr J Reprod Health 1999; 3: 51-65.

[20] United Nations Development Fund for Women (UNIFEM). Get the facts: Violence against young women and girls. United Nations publication; 2010; 1-3.