DIFFERENTIAL INCIDENCE OF DIABETES MELLITUS BETWEEN MALE AND FEMALE USING BETA-BINOMIAL AND NEGATIVE BINOMIAL MODELS IN EKITI WEST

BY

Matthew AINA

P15SCMT8059

DEPARTMENT OF STATISTICS

FACULTY OF PHYSICAL SCIENCES

AHMADU BELLO UNIVERSITY,

ZARIA, NIGERIA

JUNE, 2018

DIFFERENTIAL INCIDENCE OF DIABETES MELLITUS BETWEEN MALE AND FEMALE USING BETA-BINOMIAL AND NEGATIVE BINOMIAL MODELS IN EKITI WEST

BY

Matthew AINA
P15SCMT8059

A THESIS SUBMITTED TO THE SCHOOL OF POSTGRADUATE STUDIES, AHMADU BELLO UNIVERSITY, ZARIA-NIGERIA, IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF MASTER OF SCIENCE DEGREE IN STATISTICS

DEPARTMENT OF STATISTICS

FACULTY OF PHYSICAL SCIENCES
AHMADU BELLO UNIVERSITY,
ZARIA, NIGERIA

JUNE, 2018

Declaration

I declare that the work in this dissertation entitled "DIFFERENTIAL INCIDENCE OF DIABETES MELLITUS BETWEEN MALE AND FEMALE USING Beta-Binomial AND Negative Binomial MODELS" has been performed by me in the Department of Statistics under the supervision of Dr.T.Dahiru and Dr.H.G.Dikko. The information derived from the literature has been acknowledged in the text and a list of references provided. No part of this dissertation was previously presented for another degree or diploma at this or any other institution.

.

Name of Student	Signature	Date

Certification

This dissertation entitled "differential incidence of diabetes mellitus between male and female using Beta-Binomial and Negative Binomial models" by Matthew AINA meets the regulations governing the award of the degree of Master of Science of Ahmadu Bello University, and is approved for its contribution to knowledge and literary presentation.

.

Chairman Supervisory Committee Dr. T. Dahiru	(Signature)	(Date)
Member Supervisory Committee Dr. H. G Dikko	(Signature)	(Date)
Head of Department Dr. H.G Dikko	(Signature)	(Date)
Dean, Postgraduate School Prof. S.Z Abubakar	(Signature)	(Date)

Dedication

This research work is dedicated to God Almighty who has been my sustenance since the start of this programme.

Acknowledgements

I show my sincere gratitude to God who has given me the ability to carry out my course work and research successfully. I am sincerely grateful for the intellectual guidance and supervision of my supervisors; Dr T. Dahiru and Dr. H. G Dikko who ensured the successful completion of the work in appropriate time and the entire staff of the Department of Statistics.

I must also commend my boss in my place of work Mr. S. A. Phillips for his complete support during the course of my programme. My sincere thanks goes to my beloved wife Mrs. Aina Olubunmi and my children Ayomide Aina, Semilore Aina, and Korede Aina, I pray that the height I cannot attain you will all go beyond it.

I also appreciate the invaluable cooperation of my colleagues in this programme. I thank you all for your support and is my prayer that we all achieve our ultimate goals in life. Thank you all.

Abstract

This research entails, the differential incidence of Type 2 Diabetes Mellitus between male and female using Beta-Binomial and Negative Binomial models. From the previous works done we discovered that Type 2 Diabetes Mellitus incidence is at increase level in the Region and Localities across the country.In view of this we intend to study gender sensitivity to Type 2 Diabetes Mellitus. Secondary data for a period of 10 years were collected from medical records of General Hospital, Ifaki-Ekiti. Beta-Binomial and Negative Binomial models were used to fit the Type 2 Diabetes mellitus between male and female patients and to estimate the parameter of both model on type 2 diabetes mellitus between male and female patients.The two models were compared in term of goodness of fit and best fit using AIC with appropriate confidence interval obtained for the two fitted models . The study revealed that Beta-Binomial out performed Negative Binomial model in fitting the incidence of Type 2 Diabetes mellitus. Therefore, the result shows that Beta-Binomial is a good compound distribution to fit a model on type 2 diabetes mellitus.

Table of Contents

Cover Page……..……..i

Title Page……..……ii

Declaration.....iii

Certification.....iv

Dedication.....iv

Acknowledgements.....v

Abstract.....vi

List of Figures.....x

List of Tables.....xi

List of Appendices.....xii

CHAPTER ONE.....1

INTRODUCTION.....1

1.1Background to the Study.....1

1.2 Statement of the problem.....2

1.3 Aim and Objectives of the Study.....3

1.4 Justification of the Study.....3

CHAPTER TWO 5

LITERATURE REVIEW 5

2.1 Review of Related Literature 5

2.2 Types / Classification of Diabetes Mellitus 12

2.3 PREDIABETES 14

2.4 Limitation of Negative Binomial Distribution……..15

CHAPTER THREE 16

RESEARCH METHODOLOGY 16

3.1 Introduction 16

3.2 Method of Data Collection 16

3.3 Beta-Binomial 16

3.4 Credible interval 17

3.5 Derivation of Beta-Binomial Distribution 18

3.6 Negative Binomial 20

3.7 Derivation of Mean and Variance of Negative Binomial Distribution. 21

CHAPTER FOUR **23**

RESULTS AND DISCUSSION 23

CHAPTER FIVE 28

SUMMARY, CONCLUSION AND RECOMMENDATION 28

5.1 Summary 28

5.2 Conclusion 28

5.3 Recommendation 28

5.4 Contribution to knowledge 29

5.5 Area of Further Research 29

References 30

Appendices 32

List of Figures

Figure 4.1:Distribution of Type 2Diabetes Mellitus by Gender and Age .....**Error! Bookmark not defined.**23

Figure 4.2:Plot of Beta-Binomial Model on Type 2 Diabetes Mellitus. 274

Figure 4.3:Plot of Negative Binomial Model on Type 2 Diabetes Mellitus 24

Figure 4.4: Remedy to Over-dispersion of Negative Binomial by using Gamma Distribution……25

List of Tables

Table 4.1: Descriptive Statistics on Diabetes Mellitus between Male and Female. 23

Table 4.2: Output of Beta-Binomial Distribution 23

Table 4.3:Output of Negative Binomial Distribution 23

List of Appendices

CHAPTER ONE

INTRODUCTION

1.1.Background to the Study

Diabetes mellitus is a group of diseases caused by high level of blood glucose resulting from insufficient production of insulin and how insulin works. People with diabetes may develop serious complication such as heart disease, stroke, kidney failure, blindness and premature death. The term diabetes mellitus describes a metabolic disorder of multiple aetiology characterized by chronic hyperglycemia with disturbance of carbohydrate, fat and protein metabolism resulting from defecting insulin secretion and insulin action.

The incidence(rate) of diabetes mellitus has been a global, regional as well as local public health issue. According to, WHO(2016), an estimated 422 million adults were living with diabetes mellitus in 2014 and IDF(2015), showed that 415 million people were living with diabetes in which Europe accounted for 59.8 million , North America and Caribbean accounted for 44.3 million, Middle East and North Africa accounted for 35.4 million, South East Asia accounted for 78.3 million, Western Pacific accounted for 153.2 million, Africa accounted 14.2 million and South and Central America accounted for 29.6 million people with diabetes in 2015 respectively. It is also globally projected that by 2040, 642 million people will be living with diabetes (IDF,2015). In Nigeria, there were more than 1.56 million cases of diabetes in 2015, between the ages of 20 to 79 years (IDF,2015), More so, WHO,(2016) estimated that 1.7 million Nigerians were living with diabetes.

There are risk factors causing type 2 diabetes mellitus and they include overweight, unhealthy diet, physical inactivity, increasing age, high blood pressure, impaired glucose tolerance and poor nutrition during pregnancy (IDF,2015).The other risk

factors associated with Type 2 Diabetes Mellitus, is smoking, sedentary lifestyle, obesity, alcohol, consumption, and obstructive sleep, and WHO,(2016), noted that overweight and obesity are the strongest risk factors of type 2 diabetes mellitus.

Globally, there are major intervention set to prevent incidence(rate)of Type 2 Diabetes Mellitus. The interventions include dietary intervention or physical activity intervention or both. It is discovered that combining dietary intervention and physical activity intervention is considered to be effective and efficient, According to (ADA,2007), an individual intervention is also effective by changing of lifestyle and avoidance of all the risk factors associated with Type 2 diabetes mellitus.

The effect of diabetes mellitus include long-term damage, dysfunction and failure of various organs. It may present with characteristics symptoms such as thirst, polyuria, blurring of vision and weight loss. In most severe forms ketoacidosis or a non-ketotic hyperosmolar(shortage of insulin, in response the body switches to burning fatty acid which produces acidic ketone) may develop, frequent sweating which can lead to coma and in the absence of effective treatment, death may result.

1.2 Statement of the problem

The issue of type 2 diabetes mellitus to every state in the Nation has become problematic or order of the day and demand urgent attention from the government,to solve the incidence of type 2 diabetes while there is growing studies on diabetes.

In a different study, Chukwu,(2014), found out that there is high incidence of diabetes in the rural area than the urban area in Udi local Government Area of Enugu State.

Lawrence *et al.*, (2013), use Bayesian small area estimates of diabetes mellitus incidence, they find out that there is high level of diabetes in small countries in U.S.A.

Ebenezer *et al.*, (2003), studied the incidence (rate) of Type 2 Diabetes Mellitus(when the blood glucose level rise above 7.0mmol/L) in Port-Harcourt. He discovered that diabetes is on high increase in Port-Harcourt.

In this research, we want to compare two models to know which distribution will best fit gender sensitivity (prone to disease) to Type 2 Diabetes Mellitus based on Beta-Binomial and Negative Binomial models, we intend to check the blood glucose level of the patients from their medical records.

Therefore, in line with the study this research wish to examine Beta-Binomial and Negative Binomial models to determine which model best describe gender sensitivity to type 2 diabetes mellitus.

1.3 Aim and Objectives of the Study

The aim of this research work is to investigate the differential incidence of type 2 diabetes mellitus between male and female patients using Beta-Binomial and Negative Binomial models

The specific objectives are to:

i. Estimate the parameter of the two models on Type 2 Diabetes Mellitus between male and female.
ii. Compare between the two models usingAkaike Information Criterion (AIC) to determine the best fitted model; and
iii. Compare the confidence interval between the two fitted models.

1.4 Justification of the Study

The nature of this research will help us to determine which models will best describe the gender sensitivity to type 2 diabetes mellitus out of both beta-binomial and negative binomial distribution, since most of the literature study aggregate incidence to types 2

diabetes mellitus by using various statistical tools like chi-square, t-test and many more to justify their claims.

Therefore, this study differs from most existing literatures on the account that the study will employ Beta-Binomial and Negative Binomial models to carry out comparison in order to estimate the probabilitiesof both models between genders.

To compare between the two model their Akaike Information Criterion, and determine the confidence interval in order to show variability between gender sensitivity to diabetes mellitus.

CHAPTER TWO

LITERATURE REVIEW

2.1 Review of Related Literature

Diabetes occurs when the body becomes unable to handle glucose which builds up to dangerous levels in the blood. A diagnosis of the disease is usually made when a blood sugar test is consistently above (7.0 mmo1/L) after an eight hour fast. Fasting blood sugar levels of (6.1-6.9mmo1/L) are known as impaired fasting glucose tolerance, and usually precede full– blown diabetes. Diabetes Mellitus is a group diseases that result in too much sugar in blood (high blood glucose)

Victor (2017), used Negative Binomial model to determine the null hypothesis, to show the proportion of successes in each performed experiment and also the number of failure is larger. The experiment shows that Poisson distribution is a good approximation for the negative binomial model and it is also established that the negative Binomial model can be used to calculate the confidence interval for percentage in terms of the generalized hypergeometric function.The author recommended the areas of application of Negative binomial model such as health care, quality control, science and social science. It was further stressed that Negative Binomial model can be used in Aging research in clinical epidemiology, thunderstorm parameter of occurrence in meteorology, stock control problems and accident statistics as good examples among other applications of Negative Binomial model.

John *et al*., (2016), used Negative Binomial model to identify whether changes in overall sales are accounted for by previous non-buyers, light buyers or heavy buyers to industrial purchasing situation. The result displayed that Negative Binomial model is a good model that can be used to show consumer behavior pattern in purchasingand also the Negative Binomial model has been widely used in consumer marketing for such

competitive analysis. The result also shows that some business markets may have a heavy concentration of the light buyers, and the collective potential of these light buyers may exceed that of the few heavy buyers. Such a situation implies that these light buyers are a key to improving market share, consistent with findings in consumer research out to newer light buyer. And it is recommended that the use of the Negative Binomial distribution model would enable identification and real changes in long-term purchasing rates, perhaps due to successful marketing activities such as advertising or sales promotion. It is also shown that the Beta-Binomial model is a reasonable choice for modeling purchasing for a particular brand of buyers.

According to AJDM (2016), diabetes mellitus is a multi-factorial disease that requires long – term care since it involves major changes in both physical and psychosocial dimension of each patient. Diabetes education is a critical element of care that improves patient outcomes.

Ani *et al.*,(2016), used Bayesian Mixture Model Averaging (BMMA) as a Bayesian solution to form the best single model from all the possible mixture model,BMMA is a combination of the two methods, the Bayesian Mixture Model and Bayesian Model Averaging. The modeling of BMM is for Diabetes Mellitus data, while Bayesian Model Averaging is used for microarray data. The BMMA is used to model blood sugar levels of Diabetes Mellitus patient through simulation studies. The results of the study succeeded in building the BMMA normal model with two components mixture that could accommodate the real condition of the diabetes mellitus data with driven data concept. They opined that the model is capable of accommodating the real condition of Diabetes mellitus driven data concept in the nearest future.

Gregory and Marc (2016), used Differential-Express Genes (DEGs) of two or more conditions the results show that DEGs overestimating dispersion and therefore variance is moderately high to genes, but the result of the Beta-Binomial description can correct this to achieve better sensitivity for medium to highly express genes, and therefore, recommended that the study of mRNAseq process is inherently a binomial process and Beta-Binomial model is an appropriate choice for describing mRNA data.

Emenyong *et al.,* (2015) applied Bayesian Logistic Regression Model (BLM) to model the risk factors associated with Diabetes Mellitus. The results were compared for Bayesian logistic regression model with those from the Frequentist Logistic Regression Model (FLR) based on the significant factors. It showed that BLM with non – informative flat prior distribution and frequentist logistic regression model yielded similar results.

Jennifer *et al.,* (2014), used Negative Binomial to examine data on intravenous catheter IV insertion on sample of children. The goal of the study was to determine the appropriateness use of Negative Binomial to examine data in southern U.S hospital. The assumptions governing Negative Binomial model was examined, that is may be the data violate the assumption of independence. That result show that negative binominalis a good model fit for modeling the catheter IV insertion process data, with Akaike Information Criterion(AIC) of the Negative Binomial model indicates that it is a better fit.The result shows Negative Binomial model that appears to give more precise effects.

Anthonia and Chukwuna,2014, In the early nineties, not much was known about diabetes mellitus in Nigeria, and traditionally people related diabetes mellitus to curse or hexes and diagnosis was made based on blood or urinary test for glucose.

Tomi *et al.,* (2014), used projection predictive covariate selection to model event risk in Diabetic individuals with measurement result that were available. The finding showed that estimates of predictive performance were computed using cross validating. They recommended that a comparing of the horse-shoe, Laplace and Gaussian prior on the candidate biomarker effect demonstrated empirically and expected difference in behaviour.

Chukwu (2014), used descriptive survey research design for the examine to study the incidence and demographic differentials among Diabetes Mellitus patients by using Multi-Cluster sampling make inference. The result showed that incidence of Type 2 diabetes mellitus in 2008 and 2012 were 2.9% and 30.5% respectively. He recommended that sensitization of the general public during antenatal clinic, intensive health education program should be put in place in mass media to educate masses on signs/symptoms and consequence of diabetes mellitus.

Seam *et al.,* (2013) used new statistical approach, a physiologically based model of blood glucose dynamic; a simulation method was conducted by using Markov Chain Monte Carlo method. The finding showed that predictable performance was unaffected by parameters misspecification and was insensitive to misleading prior distributions, and the proposed methods represented a new paradigm for analysis of deterministic Mathematics model of blood glucose concentration. They suggested that future work may investigate the potential association of parameters estimates and physical characteristic (percentage of body fat, cardiovascular fitness).

Luis and Ernesto (2013), used Bayesian surprise and optimal action selection and discovered that based on Bayesian surprise and optimal action selection, a novel framework for control loop for monitoring the presence of diabetes. They observed that a unique property of the proposed approach was that it could be used for

automated supervision of Real – Time Optimization (RTO) schemes working in uncertain environments.

Atanu (2013) used the joint longitudinal modeling to compare the drug treatment effect(i.e"*Metformin plus pioglitazone*" and "*Gticlazide plus pioglitazone*") in presence of repeatedly measured missing observations to control serum cretinine levels in type 2 Diabetes patients. The Markov Chain Monte Carlo (MCMC) was used to carry out the iteration procedure. They found out that "metformin plus pioglitazone" was found more effective to control serum creatine in comparison to "Gticlazide plus pioglitazone". They recommended that elaborate method could be applied in other clinical problems to reduce the inconsistency due to presence of missing observations.

Ezebuiro *et al.*,(2013),posited that people with diabetes experience sexual weakness, delivery of large babies,still birth or recurrent abortion,becoming tired /weak/fatigued despite adequate and balanced food intake and when an individual had not done any strenuous task. Weakness in diabetes results from non-utilization of glucose by the organs of the body. In a like manner, diabetes is an immunosuppressive disorder and it is all due to hyperglycemia's toxicity in the immune system. Similarly, frequent skin infection is another symptom of diabetes and it even becomes more common when the blood sugar rises so much. In addition, wounds that are slow to heal particularly sores on the feet and ankles, and recurrent virginal infection (in women), are also signs of possible diabetes that are somewhat more likely to draw attention and result in diagnosis.

According to Marina and Maria (2012), Diabetes Mellitus (DM) is a diverse group of metabolic disorder that is often associated with a high disease burden in developing countries such as Nigeria.

Lizhen *et al.*, (2012), used Markov Chain Monte Carlos (MCMC), and found out that the postulated latent variable represents the underlying severity or complication level of trait and characterized the totality of multiple observed phenotype of interest and recommended that the uncertainty inherited from the first stage selection, must be accounted for in the model used in the second stage.

Fredrik (2012), used linear and Bayesian ensemble modeling to investigate simplified linear and merged models of glucose dynamics for the purpose of short term prediction. The results showed that diabetes glucose dynamic is known to comprise both short and more long- term time variability. Merging different diversified models may prove to be a successful approach, as a means to improve performance and robustness under such condition and recommended that the model could be used, in a decision support system, to alert the user of future low and high glucose levels and then implemented in a control framework to suggest proactive actions.

Parthiban *et al.*,(2011) used naïve Bayesian data mining classifier technique. The finding shows clearly that the proposed methods performs well compared to other similar methods in the literature taking into the fact that the attributes taken for analysis were not direct indication of heart disease which were not direct indication of the class we were trying to predict. They recommend that the data analysis result could be used for further research in enhancing the accuracy of the prediction system in future.

Yi *et.al*,(2010), used ordinary differential equation and found out that the glucose – insulin minimal model provided the result which indicated that the Bayesian method applied a satisfactory solution with accurate parameter estimate which were numerically stable.They recommended that Bayesian approach should be considered

as an additional tool for data analysis, since it enables the effect modelling and provides precise parameter estimate.

Lawrence and Thompson (2009), used diffuse prior or uninformative prior to estimate the incidence of Diabetes and prevalence of the disease in the Counties. The result showed that there was high incidence of diabetes in the Southern Eastern U.S, the Appalachian region and scattered counties throughout the Western U.S.

Thang *et al.*,(2009), used Fisher's Exact test, G-test, t-test and Local pooled-error techniques for differential analysis of spectra count data in two cancer experiment, the result showed that the method was unable to declare 95% confidence interval level, and also the result had a short coming that they could not take into account within and between sample variation together. But Beta-Binomial model incorporates both the within and between sample variation, which naturally normalizes for total count. The Beta-Binomial test performs favorably in comparison with other methods on several datasets in term of both true detection rate and false positive rate. They recommended that Beta-Binomial model be applied for experiment with one or more replicates and multiple condition comparison

Shakil (2005),used Beta-Binomial model to analyze questions of the state examination, and used the distribution to assess student performance based on the questions in the state examination. The result created a valid and reliable inference and recommended that the model be used to describe the classroom test and the use to test anxiety, attitude survey and general learning outcome.

Ebenezer *et al.*, (2003),used two stage cluster sampling technique to conduct survey on diabetic subjects to measure the plasma glucose level and also consider the associated risk factors of type 2 Diabetes Mellitus. The result showed that the prevalence of type2 Diabetes mellitus in Port- Harcourt is relatively high and

recommended that type 2 Diabetes to a large extent was a preventable disease which could be prevented by change of life style.

Michael(2003),used Beta-Binomial distribution to estimate the matching performance of a biometric identification. The Beta-Binomial model was used to assess the variability in estimates of the false match and the false non-match rate when multiple users are tested more than once. The result account for the extraneous variability and allows for the creation of confidence interval under certain regularity condition. Also, the Beta-Binomial has flexibility to model the correlation of observations which the Binomial does not possess.

Jack and Darius (1987), used Bayesian procedure for the Beta-Binomial Model and using a suitable re-parameterization,the simulation result effectively demonstrated the superiority of Bayesian estimator particularly in small samples even with uniform ("non-informed") priors. Naturally informed prior can give better result.

2.2 Types / Classification of Diabetes Mellitus

We have several types of diabetes mellitus such as:

2.2.1. Type 1 diabetes mellitus:

The Type 1 diabetes mellitus was previously called insulin dependent diabetes mellitus or juvenile- onset diabetes. Although disease onset could occur at any age, the peak age for diagnosis is in the Mid- teens. Type 1 diabetes develops when the cells that produce the hormone insulin, known as the beta cells in the pancreas are destroyed. This destruction is initiated or mediated by the body's immune system and limits or completely eliminates the production and secretion of insulin, the hormone that is required to lower glucose levels. To survive, people with type 1 diabetes must have insulin delivered by injection into their body system.

2.2.2. Type 2 diabetes mellitus:

The Type 2 diabetes mellitus was previously called non-insulin dependent diabetes mellitus or adult – onset diabetes because the peak age of onset is usually later than type 1 diabetes. In adult, type 2 diabetes usually begins with insulin resistance, a disorder in which the cells primarily within the muscles, liver and fat tissue do not use insulin properly. As the need for insulin rises, the beta cell in the pancreas gradually lose the ability to produce sufficient quantities of the hormone. Furthermore, Type 2 diabetes impact men and women proportionately. There are over 12 million men with diabetes and 11.5 million women with diabetes(flushmind,2016H). The blood sugar concentration or blood glucose level is the amount of glucose (sugar) present in the blood of man/woman. Normal blood sugar level in the blood enhances the capacity of an individual to function effectively and efficiently in his/her daily activities. A fasting glucose level of 7.0mmo1/L denote pre-diabetes. The normal, non-diabetes adult's human body maintains the blood sugar level at a range between 6.1-6.9 mmo1/L and rise to nearly or 7.0mmo1/L a bit moved after a full meal. Fasting blood sugar level outside the normal range that rises to 7.0mmo1/L or more is an indicator of medical condition clinically diagnosed as hyperglycemia (Chukwu,2014).

2.2.3. Gestational diabetes:

Is a form of glucose intolerance diagnosed during the second or trimester of pregnancy. During pregnancy, increasing blood glucose level increase the risk for both mother and fetus and requires treatment to reduce problems for the mother and the infant. Treatment for such includes special diet, regular physical activity or insulin injection.

2.2.4Maturity Onset Diabetes of the Young (MODY):

Is a rare form of diabetes which is different from both type 1 and type 2 diabetes mellitus and runs strongly in families. MODY is caused by a mutation (or change) in a single gene. It affects 1-2% of people with diabetes although it often goes unrecognized.

The 3 main features of MODY are:

1. Diabetes often develop before the age of 25.
2. Diabetes runs in families from one generation to the next.
3. Diabetes may be treated by diet or tablet and does not always need insulin treatment.

2.2.5. Latent Autoimmune Diabetes in Adults (LADA):

Is a disorder in which, despite the presence of islet antibodies at diagnosis of diabetes, the progression of autoimmune β cell failure is slow. LADA patients are therefore not insulin requiring, at least during the first 6 months after diagnosis of diabetes mellitus(ADA,2005).

2.3 Prediabetes

Pre-diabetes is a condition in which individuals have high blood glucose levels but not high enough to be classified as diabetic. People with pre-diabetes have an increased risk of developing type 2 diabetes, heart disease, and stroke, but not everyone with pre diabetes will progress to diabetes. WHO, (2006), Diabetes is on the rise. It is no longer a disease of predominantly rich nations, the prevalence of diabetes is steadily increasing everywhere, most markedly in the world of middle income countries. When diabetes is uncontrolled, it has dire consequences on health and well-being.The results of our findings support the work of work of (Micheal,2003) and (Shakil,2005) which showed that Beta-Binomial Model is a good modelin conducting an experiments and the result of John *et al.*,(2016),shows

that Negative Binomial is a good model to compare experiments but it results were always showing overdispersion as compared to that of Beta Binomial of lower dispersion, this shows that Beta Binomial can outperform Negative Binomial in the same given experiments.

2.4 Limitation of Negative Binomial Distribution

From our results it was discovered that Negative Binomial show high level of over dispersion in it results.This calls for the urgent remedy to solve the issue of overdispersion.In order,to solve this problem we used Gamma distribution to control the high dispersed values,although we have other approach like the Chi Square distribution but we used Gamma distribution to control the dispersion.We found out that the values of the standard error and confidence interval was brought down to a low significant level.

CHAPTER THREE
RESEARCH METHODOLOGY

3.1 Introduction

In this research work, Beta-Binomial and Negative Binomial models will be used to carry out the research. Beta-Binomial is a compound distribution which comprises two distributions which are Beta distribution as a continuous distribution and Binomial as a discrete distribution, which are mostly used in a Bayesian predictive model and can also be used in various aspects of inferential statistics research methods ranging from predictive, comparison, and evaluation of experiments and Negative Binomial model can be used in term of counting experiment and also to determine success and failure in an experiment ,just like the Beta-binomial model it can also be used for comparison ,predictive and evaluation of experiments.

3.2 Method of Data Collection

The data we are considering is a secondary data which will be retrieved from patient medical records for a period of 10 years at the General Hospital, Ifaki-Ekiti, Ekiti-State. From each patient's files we shall look into individual Glucose level and record for both male and female patients. This will enable us to determine whether patients are pre-diabetic, if the blood glucose level rises to nearly 7.0mmo1/L and any patient whose blood glucose level rises above 7.0mmo1/L is an indicator of Type 2 Diabetes Mellitus in patients. The data used for this research work is presented in the appendix pages.

3.3 Beta-Binomial

Suppose the prior distribution is Beta (α_1, α_2) and the conditional distribution of x given p is $Bin(n, p)$ then,

$$f(x \mid p) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0,1,...n \qquad 3.1$$

wherep is Beta distribution with parameters$\alpha_1$and α_2i. e β~ Beta(α1, α2).

Posterior is Beta – Binomial with the parameters α1 and α2.

where x is binomial ~ bin(n,p).

$f(x \mid p)$is the conditional distribution of the occurrence of diabetes given the parameter of the occurrence of the disease which follow binomial distribution.

$$g(p) = \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} p^{\alpha_1 - 1}(1-p)^{\alpha_2 - 1} \quad , 0 < p < 1 \qquad 3.2$$

$g(p)$ is parameter density function(PDF) of Beta distribution.

Therefore, the compound distribution is the Beta-Binomial given below which is discrete in nature,due to the combination of both binomial and Beta distributions.The probability mass function(PMF) is given as follows:

$$f(x \mid p)g(p) = \binom{n}{x} \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} p^{x+\alpha_1 - 1}(1-p)^{n-x+\alpha_2 - 1} \qquad 3.3$$

Suppose the data x are Bin (n, p) and the prior is Beta (α_1, α_2), so the posterior is Beta $(x+\alpha_1; n-x+\alpha_2)$

3.4 Akaike Information Criterion

Akaike Information Criterion is usually calculated in the software.The basic formula is defined as:

AIC= -2(Log-likelihood) + 2k

where

- K is the number of model parameters.
- Log-Likelihood is a measure of model fit .The higher the number ,the better the fit.

3.5 Credible interval

One way to make credible intervals is to find marginal posterior distribution for the parameter of interest and find it's $\frac{\alpha}{2}$, the interval between them is called credible interval.

In this research, we shall adopt the use of Beta-Binomial and Negative Binomial distribution.

Another name for credible interval is called the highest posterior density region.

3.6 Derivation of Beta Binomial Distribution

Suppose that there are M individual test items and each of those individual test items is tested n times. Let be X is the number of success, and

$$P(X = x_i) = \binom{n}{x_i} p^{x_i}(1-p)^{n-x_i}, \; i = 1,2,3,....n \qquad 3.4$$

$$f(p \mid \alpha_1 \alpha_2) = \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} p^{\alpha_1 - 1}(1-p)^{\alpha_2 - 1}, \; 0 \le p \le 1, \alpha_1 > 0, \alpha_2 > 0 \qquad 3.5$$

Derivation of mean of Beta-Binomial

$Let X = x_i \,/\, p, n \sim Bin(n, p)$

$$= \binom{n}{x_i} p^{x_i}(1-p)^{n-x_i} \qquad 3.6$$

Such that

$X E_{x_i}[X = x_i] \rightarrow mean$

$E[X = 1 \setminus p, n] = np$ mean of binomial variable

$$E[X = x_i] = \sum x_i p(X = x_i) \qquad 3.7$$

$$= \sum x_1 \int p(X = x_i, p) dp$$

$$= \sum x_1 \int p(X = x_i, \setminus p, n) f(p, n) dp$$

$$= \int f(p;n) \sum x_i p(X = x_i / p, n) dp$$

$$E[X = x_i] = \int npf(n, p) dp$$

$$E[X = x_i] = \int npBeta(p, \alpha_1 \alpha_2) dp \qquad 3.8$$

$$f(p, n) = Beta(p, \alpha_1, \alpha_2, n)$$

$$E[X] = n \int pBeta(p; \alpha_1, \alpha_2) dp$$

$$E[X] = \frac{n\alpha}{\alpha_1 + \alpha_2} \text{ mean}$$

Derivation of variance of Beta-Binomial

We will derive the second moment

$$E[X^2 / p, n] = np(1-p) + n^2 p^2$$

$$Var[X] = E[X^2] - (E(X)^2$$

$$E[X] = \frac{n\alpha_1}{(\alpha_1 + \alpha_2)} \text{ mean}$$

$$E[X^2] = \sum x_i^2 p(X = x_i) \text{ density function}$$

The product of binomial and beta binomial

$$= \sum_{x_i} x^2 \int p(X = x_i \setminus p, n).f(p; n, \alpha_1, \alpha_2) dp$$

$$= \int f(p; n, \alpha_1, \alpha_2) \sum_{x_1} x^2 p(x = 1 / p, n) dp$$

$$= \int (np(1-p) + n^2 p^2) Beta(p; \alpha_1 \alpha_2) dp \qquad 3.9$$

$$E[X^2] = nE[p] + (n^2 - n)E[P^2]$$

$$E[X^2] = \frac{n\alpha_1}{\alpha_1 + \alpha_2} + [n^2 - n[\frac{\alpha_1 \alpha_2}{(\alpha_1 + \alpha_2)(\alpha_1 + \alpha_2 + 1)} + \frac{\alpha^2}{(\alpha_1 + \alpha_2)}]$$

$$Var[X] = E[X^2] - (E(X)^2$$

$$= \frac{n\alpha_1}{\alpha_1+\alpha_2} + \frac{(n^2-n)\alpha_1\alpha_2}{(\alpha_1+\alpha_2)(\alpha_1+\alpha_2+1)} + \frac{n^2{\alpha_1}^2}{(\alpha_1+\alpha_2)^2} - \frac{n\alpha^2}{(\alpha_1+\alpha_2)} - \frac{n^2{\alpha_1}^2}{(\alpha_1+\alpha_2)^2}$$

$$\therefore Var[X] = \frac{n\alpha_1\alpha_2(\alpha_1+\alpha_2+n)}{(\alpha_1+\alpha_2)^2}$$

$$Var[X] = \frac{n\alpha_1\alpha_2(\alpha_1+\alpha_2+n)}{(\alpha_1+\alpha_2)^2(\alpha_1+\alpha_2+1)}$$

The Beta-Binomial distribution is also known as an extra variation model, because it allows for greater variability among the $x_i's$ than the binomial distribution.

3.6 Negative Binomial

The study byMichelle *et al.,(*2017) presented the description of negative Binomial as thesame Pascal distribution. The negative binomial distribution with parameter r and p denoted as the possible values of y are the nonnegative integers: 0,1,2,3……

Negative binomial regression is a generalization of Poisson regression which loosens the restrictive assumption that the variance is equal to the mean made by the position model. The traditional negative binomial regression model, commonly known as NB2, is based on the Poisson gamma mixture distribution. This formulation is popular because it allows the modeling of Poisson heterogeneity using a gamma distribution.

The parameter mass function of the negative binomial distribution is written as follows:

$$f(x;r,p) = p_r(X=x) = \binom{x+r-1}{x} p^x(1-p)^r, x = 0,1,2..... \qquad 3.13$$

where x is the number of success, r is the number failure, and p is the parameter of success.

3.7 Derivation of Mean and Variance of Negative Binomial Distribution.

The derivation of mean is as follows.

$$E[X]=\sum_i f(x_i).x_i=\sum_{x=0}^{\infty}\binom{x+r-1}{r-1}p^x(1-p)^r.x \tag{3.14}$$

$$E[X]=\binom{0+r-1}{r-1}p^0(1-p)^r.0+\sum_{x=1}^{\infty}\binom{x+r-1}{r-1}p^x(1-p)^r.x \tag{3.15}$$

$$E[X]=\frac{rp}{1-p}\sum_{x=1}^{\infty}\frac{(x+r-1)!}{r!(x-1)!}p^{x-1}(1-p)^{r+1} \tag{3.16}$$

Let $s=r+1$ and $w=X-1$ inside the summation

$$E[X]=\frac{rp}{1-p}\sum_{w=1}^{\infty}\frac{(w+s-1)!}{(s-1)!w!}p^w(1-p)^s \tag{3.17}$$

$$E[X]=\frac{rp}{1-p}\sum_{w=0}^{\infty}\binom{w+s-1}{s-1}p^w(1-p)^s \tag{3.18}$$

The summation which is the sum over the complete parameter mass function (PMF) of a negative binomial random variable distribution is 1.

$$E[X]=\frac{rp}{1-p} \tag{3.19}$$

Then the derivation of the variance can be expressed as follows:

We will calculate $E[X^2]$ and then return variance formula

$$E[X^2]=\sum_i f(x_i).x^2=\sum_{x=0}^{\infty}\binom{x+r-1}{r-1}p^x(1-p)^r.x^2 \tag{3.20}$$

$$E[X^2]=0+\sum_{x=0}^{\infty}\binom{x+r-1}{r-1}p^x(1-p)^r.x^2$$

$$E[X^2]=\sum_{x=1}^{\infty}\frac{(x+r-1)!}{(r-1)!x!}p^x(1-p)^r.x^2$$

$$E[X^2] = \frac{rp}{1-p}\sum_{x=1}^{\infty}\frac{(x+r-1)!}{r!(x-1)!}p^{x-1}(1-p)^{r+1}x$$

Let $s = r+1$ and $w = X-1$

$$E[X^2] = \frac{rp}{1-p}\sum_{w=0}^{\infty}\frac{(w+s-1)!}{(s-1)!w!}p^{w}(1-p)^{s}(w+1)$$

$$E[X^2] = \frac{rp}{1-p}\sum_{w=0}^{\infty}\binom{w+s-1}{s-1}p^{w}(1-p)^{s}w + \sum_{w=0}^{\infty}\binom{w+s-1}{s-1}p^{w}(1-p)^{s}$$

The first summation is the negative binomial random variable distributed and the second summation is the complete sum of that variable's parameter mass function (PMF) .

$$E[X^2] = \frac{rp}{1-p}\left(\frac{sp}{1-p}+1\right) \qquad 3.26$$

$$E[X^2] = \frac{rp(1+rp)}{(1-p)^2} \qquad 3.27$$

We now insert into the original variance formula.

$$Var[X] = \frac{rp(1+p)}{(1-p)^2} - \left(\frac{rp}{1-p}\right)^2 \qquad 3.28$$

$$Var[X] = \frac{rp}{(1-p)^2}$$

The negative binomial distribution is a discrete parameter distribution of the number of successes in a sequence of independent and identically distributed Bernoulli trials before a specified (non-random) number of failures(r) occurs.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 Introduction

This chapter presents an overview of the analysis of the study and its findings. The results will show the differential analysis of incidence of type 2 diabetes mellitups between male and female by using beta binomial and negative binomial models to established which of these models is best and most efficient to fit differential incidence of Type 2 Diabetes mellitus between Male and Female patients.

4.2. Descriptive Statistic

Table 4.1: Descriptive Statistics on Diabetes Mellitus between Male and Female Patient.

	n	TOTAL(N)	PERCENTAGE
MALE	1360	6400	21.25
FEMALE	840	5650	14.87
	2200	12050	

Comment [S1]: This is not the mean of gender, they are just the sum of gender present in your data.

Table 4.1 present the descriptive statistics of gender sensitivity to type 2 diabetes mellitus. The table shows that 1360 males were affected with type 2 diabetes mellitus between years under review out of 6400 male patients reported and corresponding total of 840 females were equally affected with the diseases out of 5650 female patients reported cases. This means that more male were sensitive to the diseases which accounted for 21.25% of the total male affected with corresponding total females affected to be 14.87% respectively. Out of the total number of patients (12050) reported in the hospital for the period under review 18.26% accounted for the incidence cases of Type 2 Diabetes Mellitus.

Table 4.2: Output of Beta-Binomial Model.

GENDER	PARAMETERS	LL	AIC	STD.ERR	95% C.I
MALE(α_1)	0.9750	-2.7984	7.2968	0.2838	0.4188-1.5312
FEMALE(α_2)	0.7892	-2.8037	7.6074	0.0184	0.7479-0.8201

Table 4.3:Output of Negative Binomial Model.

GENDER	PARAMETERS	LL	AIC	STD.ERR	95% C.I
MALE (p_1)	0.7984	-67.65	137	126.15	-246.47-248.05
FEMALE(p_2)	0.7440	-106.89	216	87	-169.77-171.26

Table 4.2 and 4.3, show that two models Beta-Binomial and Negative Binomial models were used to fit type 2 diabetes mellitus on the basis of gender patients (Male and Female patients) respectively.Table 4.2, show the parameter of 0.975 while table 4.3, show the parameter 0.7984 for both Beta-Binomial model and Negative Binomial model respectively, this means that Beta-Binomial model have high chances to best fit the data than the Negative Binomial model.This indicate that Beta-Binomial model is relatively good to fit diabetic model than Negative Binomial model. More so, the values of the 95% confidence interval were extremely large for the Negative Binomial unlike the small values of Beta-Binomial model,which also proved that Beta-Binomial is a better fitted model.

In order to show the validity of our two models we compare the values of there AIC for both models. The results show that the AIC of Beta-Binomial is 7.2968 for male patients and Negative Binomial model is 137.This shows that Beta-Binomial model is better and efficient to fit Male type 2 diabetes patients than Negative binomial models.

In addition, when we compare both Table 4.2 and 4.3 for female diabetes patients. The parameter of female affected with type 2 diabetes mellitus is 0.7892 for Beta-Binomial, while the Negative Binomial modelparameter is 0.7440 which is lower than Beta-Binomial model.This indicates that Beta-Binomial model have high chances to fit type 2 diabetes model than Negative Binomial model.The 95%

confidence interval for Beta-Binomial model is between 0.7479-0.8201, compared to lower and upper bound of Negative Binomial model which ranges between -169.77-171.26 which is too extreme compare to Beta-Binomial model.This imply that Beta-Binomial model is a good model to fit female type 2 diabetes mellitus than the Negative Binomial model.

Finally, the AIC for Beta Binomial is 7.6074 for female as compared to AIC of Negative Binomial of 216 for female, and the model with lowest AIC is better than AIC with higher values, means that Beta Binomial model perform better than negative binomial model.

Based on the various results displayed in both tables 4.2 and 4.3 respectively, it shows that Beta Binomial is the best model to fit both male and female patient's type 2 diabetes model. When we further compared the value of AICof male and female patients in Table 4.2 and 4.3 respectively, the value of AIC of Beta-Binomial distribution is lowerthan the value of AIC of Negative Binomial distribution. This means that Beta-Binomial is better model than Negative Binomial model to fit both male and female distribution model. More so, in Table 4.2, the range of dispersion is extremely small compared to the value of Negative Binomial distribution which has the extreme values in both lower and upper region of Table 4.3.This shows that Beta-Binomial is a good model to fit the Diabetes mellitus than Negative Binomial distribution.

TABLE 4.4: Remedy To Over dispersion of Negative Binomial Distribution By Using Gamma Distribution.

GENDER	PARAMETER	LL	AIC	STD.ERR	95% C.I
Male(γ_1)	0.8180	-67.66	139.32	0.2729	0.2831-1.3529
Female(γ_2)	2.3679	-102.75	209.51	0.6219	1.1489-3.5868

We discovered that the Gamma distribution brought down the values of both standard error and the confidence interval in table 4.4 as compared with the values of Negative Binomial in table 4.3.Therefore ,Gamma distribution can control values of over dispersion of the Negative Binomial Distribution.

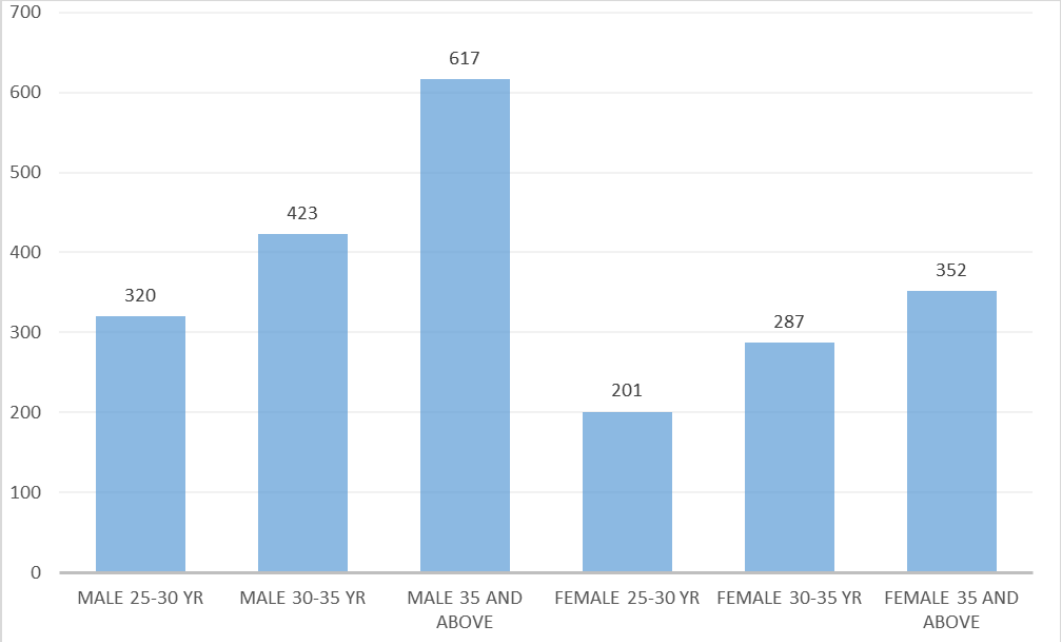


Figure 4.1:Distribution of Type 2Diabetes Mellitus by Gender and Age.

Figure 4.1**:** shows the distribution of type 2 diabetes mellitus by gender and age group.We found out that as the age increases there is high sensitivity to the disease, but male patients are more liable to the disease at age of 35 and above,than female patients in the corresponding age.We notice that as the age increases there is an increase in the number of reported cases of type 2 diabetes mellitus.

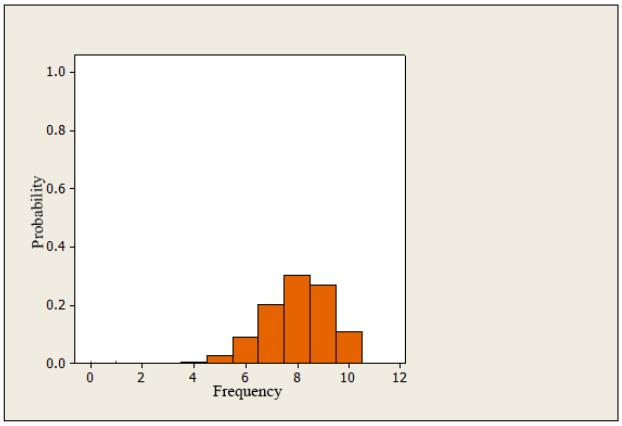


Figure 4.1:Plot of Beta-Binomial Model on Type 2 Diabetes Mellitus.

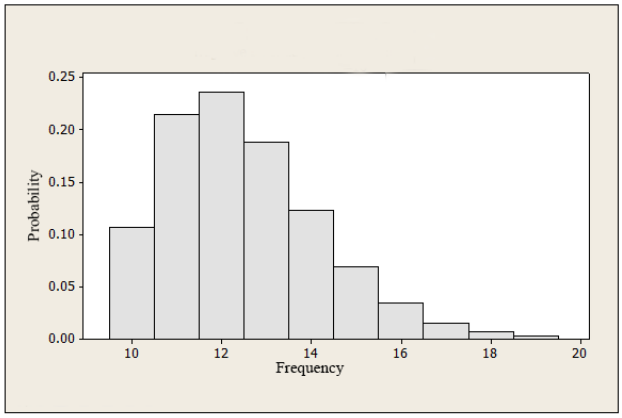


Figure 4.3:Plot of Negative Binomial Model on Type 2 Diabetes Mellitus.

Figure 4.2 and 4.3, shows the Histogram of Beta-Binomial distribution skewed to the left from the frequency 8 to the frequency 4 with parameter 0.35 while the Negative Binomial Histogram skewed to the right, from the highest frequency of 10 at the parameter 0.23. If we compare both Histogram with their parameter it shows that Beta-Binomial is better in fitting the type 2 diabetes mellitus model.

CHAPTER FIVE

SUMMARY, CONCLUSION AND RECOMMENDATION

5.1 Summary

In this research work, we investigated the distribution ofmodels that best fit the type 2 diabetes mellitus. Out of both Beta-Binomial and Negative Binomial model, we discovered that Beta-Binomial distribution is a better model than the Negative Binomial distribution .When compared their confidence interval andparameter and when we used AIC to validate our results, we discovered that Beta-Binomial model AIC model show small value compared with Negative Binomial model AIC model meaning that AIC of Beta Binomial is better than AIC of Negative Binomial model.

We found out that the confidence interval of Negative Binomial shows extreme values than the confidence interval of Beta-Binomial model with small values.

The Figure 4.2 and 4.3, shows the Histogram of Beta-Binomial distribution skewed to the left from the frequency of 8 to the frequency 4 with parameter 0.35 while the Negative Binomial model, Histogram slope to the right from the highest frequency of 10 at the parameter 0.23. If we compare both Histograms with their parameter it shows that Beta-Binomial model is better in fitting the type 2 diabetes model.

5.2 Conclusion

Based on Beta-Binomial and Negative Binomial models on comparative basis shows that Beta-Binomial model is a better model to fit Type 2 Diabetes mellitus model than Negative Binomial model as results of the various comparison outputs including AIC which is used for model selection and validation.

5.3 Recommendation

In line with our findings, it is recommended that Beta-Binomial model is a very good distribution to model gender sensitivity to Type 2 diabetes mellitus.

It is also recommended to use the Beta-Binomial in other area of experimental purposes like engineering, social science, agriculture and so on.

5.4 Contribution to knowledge

This research used both Beta-Binomial and Negative binomial to fit Male and Female sensitivity to Type 2 Diabetes Mellitus on comparative basis to determine the best fitted model. The results show that Beta-Binomial is better than Negative binomial model. In view of this Beta-Binomial is recommended an efficient model than the Negative Binomial model in the study of Type 2 diabetes mellitus.

5.5 Area of Further Research

Since our study is based on both Beta-Binomial and Negative Binomial models to know the best fitted model to fit both Male and Female type 2 Diabetes Mellitus, it is recommended for area of further research the to investigate other distributions that can perform better than Beta Binomial distribution like Poisson,Binomial, Poisson Regression.It is also recommended that this model should be used to compare with other distributions to ascertain it efficiency and validity.

References

Comment [S2]: Follow PG format of referencing. Most of your reference is not in-line. Check and do a nice work in this area.

Ani, B.A, Nur,I, Irhamah, Heri,K, and Laksmi,S. (2016).Blood sugar levels of Diabetes Mellitus Patients Modeling with Bayesian mixture Model Averaging. *Global Journal of Pure and Applied Mathematics.* 12(2): 3143-3158.

Anthonia O O, Chukwuma E. (2014). Diabetes mellitus in Nigeria: The past, present and future. *World Journal of Diabetes*, 8: 30-55.

Atanu,B. (2013).A Bayesian Joint Analysis and Imputation Model for Longitudinal: An Application in Type 2 Diabetes Drug effect Comparison. *International Journal of Collaborative Research on Internal Medicine & Public Health*.5(2):103-111.

Chukwu, B.N. (2014).Incidence and Demographic Deferential among Diabetes Mellitus Patients Attending Hospital/Clinic in Udi Local Government Area of Enugu-State.4(8):131-138.

Ebenezer, A.N, Osaretin,J.O, Anele,E.I, Aaron,O.and Seye,B .(2003).Type 2 Diabetes in adult Nigerians. A study of its prevalence and risk factors in Port-Harcourt. Diabetes Research and Clinical Practice **62**(23): 177-185.

Emenyong,S.C,Mohd,B.A, Istrrinayagy,K, Shamarina,S, and Chris,B.G.(2015). Bayesian Logistic Regression model on risk factors of Type2 Diabetes Mellitus. *Mathematical Theory and Modeling*.5(1):2224-5804.

Ezebuiro,V.O, Samuel, E.S, Nwachukwu, K.C,Bernice,C.N. (2013). Gender Differential in the incidence of Diabetes Mellitus among patients in Udi Local Government Area of Enugu State,Nigeria. *Mediterranean Journal of Social Science,* 4(8): 131-140.

Fredrik.S. (2012). Diabetes Mellitus Glucose prediction by linearand Bayesian Ensemble Modeling. Lund, Sweden: University of Lund printing Press.6(1):126-128.

Global Report on Diabetes: World Health Organization (2016).*http://www.who.int.*

Gregory, R. and Marc, R. B, (2016). A Mechanistic beta-binomial parameter model for Nrna sequence data.Department of pharmacology and systems therapeutics,Icah school of Americal.11(6):1-18.

IDF. (2015). International Diabetes Atlas. Karakas Print. Retrieved from http://www.diabetesatlas.org

Jack,C.L and Darius, J.S.(1987).Bayesian Estimation and prediction for the beta-Binomial model. *Journal of Business & Economics Statistics*, 5(3):.357-367.

Jennifer.M,Pamela.L,Jason.B. (2014). Exploring the use of Nagative Binomial Regression Modeling for Pediatric Peripheral Intravenous Catheterization. *Journal of Medical Statistics and Informatic*, 2: 360-367

John.W.W,Giang.T,Richard.L,Brown.N. (2016). Can the Nagative Binomial Distribution Predict Industrial Purchases? *Journal of Business and Industrial Marketing*, 1(31): 543-552.

Lawrence,E.B,Theodore,J.Thompson,J.T,Karen,A.K,James,P.B,Linda,S.G,Mary,and Ann, L.A .(2009). Bayesian small area estimate of Diabetes Incidence by U.S County. *Journal of Data Science* (11):249-267.

Lizhen,X, Radu,V.C, Lei,S,and Andrew,D.P.(2012).Bayesian Latent Variable Modeling of Longitudinal Family Data for genetic Pleistrop Studies.14(1):1211-1405.

Luis ,A. and Ernesto,M. (2013). Continuous Performance Monitoring Of A Closed - Loop Insulin Delivery System Using Bayesian Surprise. Universidale Federal de Santa Caterina.7(44):2175-8018.

Marina, K. and Maria, K. (2012). Education And Diabetes Mellitus. Quarterly Scientific, Online Publication by Department of Nursing A', Technological Education Institute of Anthens.8(2):654-662.

Michael, E. (2003). Using The Beta Binomial Distribution To Assess Performance of a Biometric Identification Device. *International Journal of Image and Graphics, 03*(03):1-18.

Michelle.D,Michael .J.K, and Daniel.Z. (2017). Clinical Trial Design Using A Stop Negative Binomial Distribution.Statistics and its interface.0(0000):1-10.

Parthiban, P. Rajesh, A. and Ksrivatsa.S.K.(2011). Diagnosis of Heart Disease for Diabetics Patients Using Naïve Bayes Method. *International Journal of Computer Applications* (0975-8887),24(*3*).

Richard, J.S. and Trevor, J.S. (2007). Estimation across multiple models with application to Bayesian and Software development.Diabetes Trials Unit, University of Oxford. London.17(3):245-252.

Seam, M.E, Sujit, K.S, John, J.V, Christopher, D.B ,and Andrew, J.C (2013),A Bayesian network for modeling blood glucose concentration and exercise in type1 diabetes. Southampton Statistical Sciences Research Institute.7(1):1-37.

Shakil, M. (2005). Using Beta -Binomial Distribution in Analyzing some multiple-choice Questions of the Final exams of a Math course, and its application in predicting the performance of future students. Miami.44(4):427-442.

Thang V.P, Sander, R.P, Marc,W ,and Connie, R.J(2009). On the beta binomial model for analysis of spectra count data in label-free tandem mass spectrometry-based proteomics.Oncoproteomics Laboratory,Department Medical oncology VUmc-Cancer centre Amsterdam,UV University Medical Centre.26(3): 363-369.

The African Journal of Diabetes Medicine, 2016. 24(**2**).

Tomi,p, Veikko,S, and Aki,V (2014). Hierarchical Bayesian survival analysis and projective covariate selection in cardiovascular event risk prediction.

Victor.N. (2017). Hypothesis Test and Confidence Interval for the Negative Binomial Distribution Coincidence:A case For Rare Event. *International Journal of Contemporary Mathematical Science*, 12(1): 243-253.

World Health Organization (WHO) and International Diabetes Federation (2006)

Yi, W. Kent, M .and Andrzej T.G (2010). Bayesian analysis of minimal model under theinsulin-modified IVGTT.Health.2(3):188-194..

APPENDICES

Socio demographic data of male and female patients affected with type 2 diabetes mellitus

2007

Appendix A

MONTH	AGE	25-30		30 – 35		35 and above	
	GENDER	M	F	M	F	M	F
January		2	3	3	2	4	3
February		3	1	1	-	2	2
March		-	1	2	3	3	1
April		1	1	1	1	4	-
May		-	2	2	1	3	2
June		-	2	3	2	5	1
July		-	2	1	-	1	1
August		-	-	4	-	4	-
September		1	2	2	3	4	-
October		4	1	2	2	-	3
November		1	-	1	1	4	-
December		2	-	3	-	5	2

2008

Socio demographic data of male and female patients affected with type 2 diabetes mellitus

Appendix B

MONTH	AGE	25-30		30 – 35		35 and above	
	GENDER	M	F	M	F	M	F
January		2	2	4	2	4	2
February		1	-	2	1	-	1
March		4	2	3	3	2	1
April		-	-	-	1	4	2
May		-	1	2	-	2	1
June		2	1	1	2	1	1
July		-	1	1	1	2	2
August		2	2	3	2	1	1
September		-	-	4	1	2	2
October		4	1	3	2	1	3
November		3	2	2	1	3	2
December		5	1	4	-	4	2

2009

Socio demographic data of male and female patients affected with type 2 diabetes mellitus

Appendix C

MONTH	AGE	25-30		30 – 35		35 and above	
	GENDER	M	F	M	F	M	F
January		1	3	2	2	4	3
February		3	1	1	1	2	2
March		-	1	2	3	3	1
April		1	1	1	1	4	1
May		-	2	2	1	3	3
June		2	2	3	2	5	1
July		-	2	1	1	1	1
August		3	-	4	2	4	-
September		1	2	2	3	4	-
October		3	1	2	2	3	3
November		1	-	1	1	4	3
December		2	-	3	-	4	2

2010

Socio demographic data of male and female patients affected with type 2 diabetes mellitus

Appendix D

MONTH	AGE	25-30		30 – 35		35 and above	
	GENDER	M	F	M	F	M	F
January		1	2	4	2	5	3
February		2	1	2	1	3	1
March		-	1	2	1	2	2
April		1	-	2	3	1	1
May		2	2	2	1	4	2
June		3	1	3	1	1	-
July		3	1	1	2	2	4
August		2	3	5	-	8	1
September		1	-	2	2	2	2
October		1	1	2	1	3	-
November		2	-	-	1	1	-
December		1	1	2	3	5	2

2011

Socio demographic data of male and female patients affected with type 2 diabetes mellitus

Appendix E

MONTH	AGE	25-30		30 – 35		35 and above	
	GENDER	M	F	M	F	M	F
January		1	2	4	1	3	4
February		3	1	1	-	2	2
March		-	1	2	3	3	1
April		1	1	1	1	4	-
May		-	2	2	1	3	2
June		2	1	4	3	5	2
July		-	2	1	-	1	1
August		-	-	4	-	4	-
September		1	2	2	3	4	-
October		4	1	2	2	-	3
November		2	3	2	2	5	2
December		2	1	3	2	4	2

2012

Socio demographic data of male and female patients affected with type 2 diabetes mellitus

Appendix F

MONTH	AGE	25-30		30 – 35		35 and above	
	GENDER	M	F	M	F	M	F
January		1	2	2	2	4	2
February		2	1	3	1	2	1
March		4	2	3	3	2	1
April		1	1	-	1	4	2
May		-	1	2	-	2	1
June		2	1	1	2	1	1
July		-	1	1	1	2	2
August		2	2	3	2	1	1
September		2	-	4	1	2	2
October		4	1	3	2	1	3
November		3	2	2	1	3	2
December		5	2	4	1	5	3

2013

Socio demographic data of male and female patients affected with type 2 diabetes mellitus

Appendix G

MONTH	AGE	25-30		30 – 35		35 and above	
	GENDER	M	F	M	F	M	F
January		1	2	2	2	5	3
February		2	1	1	1	2	1
March		1	1	2	3	3	1
April		-	1	1	1	4	1
May		1	2	2	1	6	3
June		2	2	3	2	5	1
July		-	2	1	1	1	2
August		3	-	4	2	4	3
September		1	2	2	3	4	-
October		3	1	2	2	3	3
November		1	1	1	1	5	3
December		2	-	3	1	4	2

2014

Socio demographic data of male and female patients affected with type 2 diabetes mellitus

Appendix H

MONTH	AGE	25-30		30 – 35		35 and above	
	GENDER	M	F	M	F	M	F
January		-	3	2	2	6	2
February		3	1	1	1	2	2
March		1	1	2	3	3	1
April		1	1	1	1	5	3
May		1	2	2	1	3	3
June		2	2	3	2	5	1
July		2	2	1	1	1	1
August		3	2	4	2	4	2
September		1	2	2	3	4	-
October		3	1	2	2	3	3
November		1	-	1	1	4	3
December		2	-	3	-	4	2

2015

Socio demographic data of male and female patients affected with type 2 diabetes mellitus

Appendix I

MONTH	AGE	25-30		30 – 35		35 and above	
	GENDER	M	F	M	F	M	F
January		3	2	2	2	7	2
February		4	3	1	1	2	2
March		2	2	2	3	3	1
April		2	1	1	1	4	1
May		1	2	2	1	3	3
June		1	1	3	2	5	1
July		4	2	1	1	1	1
August		3	2	4	2	4	3
September		1	2	2	3	4	2
October		3	1	2	2	3	3
November		1	1	1	1	4	3
December		2	3	3	-	6	4

2016

Socio demographic data of male and female patients affected with type 2 diabetes mellitus

Appendix J

MONTH	AGE	25-30		30 – 35		35 and above	
	GENDER	M	F	M	F	M	F
January		2	1	4	-	5	4
February		3	3	2	2	3	2
March		2	2	3	3	3	1
April		2	-	1	4	4	4
May		3	2	2	1	3	3
June		1	1	2	2	7	2
July		2	2	3	1	1	1
August		3	2	4	2	4	3
September		1	3	-	3	4	2
October		3	1	2	2	2	5
November		1	1	5	1	4	3
December		5	2	3	4	4	4