THRESHOLD FOR HANDLING SEVERITY OF OVERDISPERSION IN SOME COUNT DATA MODELS USING A FUZZY SET APPROACH

BY

Abidemi Damaris OYALADE
B. Sc, M.Sc, (Ibadan)
Matric No.:122693

A Thesis in the Department of Statistics,

Submitted to the Faculty of Science

in partial fulfillment of the requirements for the Degree of

DOCTOR OF PHILOSOPHY

of the

UNIVERSITY OF IBADAN

AUGUST 2023

CERTIFICATION

I certify that this work was carried out by Mrs Abidemi Damaris OYALADE, Matric. No.: 122693 in the Department of Statistics, Faculty of Science, University of Ibadan, Ibadan, Oyo State, Nigeria

…………………………………………………..

Supervisor

Oluwayemisi O. Alaba

B. Sc,, M.Sc, Ph.D (Ibadan)

Reader, Department of Statistics,

University of Ibadan, Nigeria.

DEDICATION

This thesis is dedicated to my sweetheart and husband – Pastor Gbenga Olubukoye Oyalade and my two sons Oyalade, Jeremiah Abiodun and Oyalade, Jude Ayomidipupo.

ACKNOWLEDGEMENTS

I sincerely acknowledge Almighty God for His unreserved love and mercy for me. I give all Glory to Him for seeing me through the successful completion of this research; may all Glory be to Him.

I am also indebted to my supervisor and mentor, Dr Oluwayemisi O. Alaba, and her family for their love, patience, and forbearing during this research.

My gratitude goes to my late parents, Mr Raimi Ayanleke and Mrs Khadijat Wahab for all your love, care, and effort on me; l will forever remember you. My gratitude also goes to my husband- my sweetheart for his love, support, and cooperation during the period of carrying out this research and my children (Mr Oyalade Jeremiah Abioun and Mr Oyalade Jude Ayomidipupo) for their endurance and love. You are so wonderful.

I am equally grateful to all the staff, both teaching, and non-teaching in the Department of Statistics: the Head of the Department, Prof. O.I. Osowole, Prof A. A. Sodipo, Prof G.N. Amahia, Prof O.I. Shittu, Prof O.E. Olubusoye, Dr Adedayo A. Adepoju, Dr J.F. Ojo, Prof Angela U. Chukwu, Prof K.O. Obisesan, Dr Oluwaseun Alawode, Dr Udomboso, Dr O. B. Akanbi, Dr O.S. Yaya, Dr F. Oyamakin, Dr Ogunde, Mrs V. Laoye, Dr A. T. Adeniran, Mr Ezichi, Miss Kehinde, Mr S. Aderinto, Mrs Arewa, Mrs Akinsebo and Mr Omoowo. I also appreciate all my classmates, 2007/2008 M.Sc programme. The staff of LISA office in the Department of Statistics, Mechanical Engineering Department, University of Ibadan, Federal School of Statistics Ibadan campus especially Dr K. Balogun, Mrs Obiekwe B .C.and the staff are highly appreciated.

My appreciation also goes to my siblings; l thank you all. I wish my uncle, Mr Adedokun Fatai were alive to witness this; but I pray God takes care of his family. l also appreciate my Pastor and mentor Prof O. Fagbola of the Agronomy Department University of Ibadan, Dr B. O. Onasanya of the Department of Mathematics, and all staffers in the Department. Also, l appreciate the state Coordinator, Deeper Life Campus Fellowship (DLCF), Oyo State Pastor Moranti K. and his family; all the families of DLCF, Pastor Opadokun, Bro Shola Oluwole of blessed memory and his

family, Bro Sunday Oni and the Oni's, Bro Balogun and his family, Bro Ojo Ojebisi and his family, Bro Ayinla and his family, Pastor Oduola and his family, Bro Yemi Okunlola and his family, Mr Ajobo and his family, Dr Oyefeso and his family. God will continue to take care of your family.

I am also grateful to my friends and classmates Tolu Adepoju and Gbolahan Tanimowo.

God bless you all.

Abidemi Damaris Oyalade

2023

ABSTRACT

Overdispersion, often associated with count data is difficult to handle by a single parameter regression model such as the Poisson regression model. Previous attempts to modify the Poisson regression model with additional parameters did not take cognisance of the different levels of overdispersion because there might be no need for modification at-times. Modification done without any need affects the standard error leading to wrong conclusions. Therefore, this study was aimed at determining the threshold for modification in some count data models when the problem of overdispersion is unavoidable.

Fuzzy c-partition was used to classify the degree of overdispersion severity into not severe, moderate, severe, and very severe. Membership function was constructed for each of the classes with its fuzzy dispersion percentage (d) range: 0 for not severe with $d \leq 10$, $\frac{(4d-40)}{210}$ for moderate with $10 < d \leq 40$, $d/70$ for severe with $40 < d \leq 70$ and 1 for very severe with $d > 70$. The universal set of the dispersion percentage, $D = \left(\frac{v-m}{m}\right) \times 100\%$, where v is the variance and m, the mean. Four models: Poisson (PO), Negative Binomial (NB), Com-Poisson (CP), and Generalised Poisson (GP) were used to simulate the benchmark for modification. Different random sample sizes, including $n = 20$ for small sample and $n = 5000$ for large sample were used with mean (μ) $= 0.01, 0.05, 1.00, 2.00$ and variance (σ^2) = 0.05, 0.50, 1.50, 2.50, respectively. The ratio of the residual deviance of PO (simplest model) to its degree of freedom was used to detect the presence of overdispersion in the count data. The averaging method was used to determine the threshold ($\bar{D}$). The models were validated with monthly road crashes data from the Federal Road Safety Corps in 36 states and the Federal Capital Territory of Nigeria between 2014-2018 and the Akaike Information Criteria (AIC) was used for model selection.

The threshold $\bar{D}$ for models PO, NB, CP and GP given that $n = 20$, were 24.2, 69.4, 34.8 and 32.6%; 26.6, 73.6, 26.5 and 27.1%; 23.1, 75.2, 25.1 and 37.1%; 30.4, 77.5, 54.9 and 24.5%, respectively. The highest $\bar{D}$, at different values of μ and σ^2 for PO, NB, CP and GP when $n = 20$ were 30.4, 77.5, 54.9 and 37.1%, respectively. For n= 5000, $\bar{D}$ were 27.7, 74.9, 22.1 and 28.3%; 27.6, 74.5, 22.2 and 28.9%; 27.9, 38.2, 22.2 and 29.2%; 28.2, 29.1, 22.2 and 28.3%, respectively. The highest $\bar{D}$, at different values of μ and σ^2 for PO, NB, CP and GP when $n = 5000$ were 28.2, 74.9, 22.2 and 29.2%, respectively, indicating points for modifications. The ratio of the residual deviance of PO to its degree of freedom is 42.0 flagging very severe overdispersion (95.5%) of road crashes having membership function of 1. The AIC for PO, NB, CP and GP were 8826.7, 8657.6, 2211.0 and 2205.4, respectively. This implies that GP is the best model.

The thresholds for modification of severity of overdispersion for Poisson, Negative Binomial, Com-Poisson, and Generalised Poisson models were determined. The determined thresholds could be used to minimise wrong conclusions arising from defective standard errors.

Keywords: Generalised Poisson model, Fuzzy set theory, Severity of overdispersion, overdispersion modification threshold

Word count: 493

TABLE OF CONTENTS

		Page
FRONT PAGE		
TITLE PAGE		i
CERTIFICATION		ii
DEDICATION		iii
ACKNOWLEDGEMENTS		iv
ABSTRACT		vi
TABLE OF CONTENTS		vii
LIST OF TABLES		x
LIST OF FIGURES		xii
ABBREVIATIONS		xiii
CHAPTER ONE: INTRODUCTION		
1.1	Background to the Study	1
1.2	Statement of the Problem	7
1.3	Justification	7
1.4	Motivation for the Study	8
1.5	Aim and Objectives	8
1.6	Significance of the Study	9
1.7	Organisation of the Dissertation Presented	9
CHAPTER TWO: LITERATURE REVIEW		
2.0	Chapter Overview	10
2.1	Terminologies	12
2.2	Theoretical Framework	13
2.3	Parameter Estimation	14
2.4	Maximum Estimation of Poisson Model	14
2.5	Com-Poisson Regression Model	15
2.6	Probability Density Function Com-Poisson Regression Model	15
2.7	Estimation of Com-Poisson Regression Model	16
2.8	Negative Binomial Regression Model	17

2.9 Parameter Estimation of Negative Binomial Regression 19
2.10 Generalised Poisson Regression Model 20
2.11 Parameter Estimation 22
2.12 Zero Inflated (ZI) Model 22
2.13 Zero Inflated Poisson (ZIP) Model) 22
2.14 Zero Inflated Generalised Poisson (ZIPG) Model 23
2.15 Zero Inflated Negative Binomial (ZINB) Model 23
2.16 Zero Truncated Model 24
2.17 Hurdle Models for Excess Zeros 24
2.18 Review Literature on Count data Model 25
2.19 Review Literature on Threshold Methods 33

CHAPTER THREE: RESEARCH METHODOLOGY

3.1 Poisson Regression Model 36
3.2 Fuzzy Set Approach 37
3.3 Fuzzy-C-Partition 38
3.4 Fuzzy Class Poisson Regression 39
3.5 Negative Binomial Distribution 41
3.6 Fuzzy Class Negative Binomial Regression 42
3.7 Com-Poisson Regression Model 43
3.8 Fuzzy Class Com-Poisson Regression Model 44
3.9 Generalised Poisson Model 44
3.10 Fuzzy Class Generalised Poisson Regression Model 44
3.11 Monte Carlo Design 45

CHAPTER FOUR: RESULTS AND DISCUSSION

4.0 Chapter Overview 47
4.1 Fuzzy Set Classification of Different Levels of Overdispersion 47
4.2 Averaging Method for Determination of Threshold 82
4.3 Application to Accident Data 88
4.4 Multicollinearity 89
4.5 Parameter Estimation and Statistical Inference 89
4.6 Exploratory Data Analysis 92

4.7 Model Selection 96
4.8 Discussion of the Results 101

CHAPTER FIVE: SUMMARY, CONCLUSION AND RECOMMENDATIONS

5.1 Summary 111
5.2 Conclusion 112
5.3 Recommendations 112
5.4 Contributions to Knowledge 113
5.4 Suggestion for further studies 113

REFERENCES
APPENDIX

LIST OF TABLES

Pages

Table 4.1:Fuzzy set classification of different levels of overdispersion of PO n=20, 30. 48

Table 4.2:Fuzzy set classification of different levels of overdispersion of PO n =50, 100. 50

Table 4.3:Fuzzy set classification of different levels of overdispersion of PO n=200, 300 52

Table 4.4:Fuzzy set classification of different levels of overdispersion of PO n =500, 1000. 54

Table 4.5:Fuzzy set classification of different levels of overdispersion of PO n=5000. 56

Table 4.6:Fuzzy set classification of different levels of overdispersion of NBR n=20, 30. 56

Table 4.7:Fuzzy set classification of different levels of overdispersion of NBR n =50, 100. 58

Table 4.8:Fuzzy set classification of different levels of overdispersion of NBR n =200, 300. 60

Table 4.9:Fuzzy set classification of different levels of overdispersion of NBR n =500, 1000 62

Table 4.10:Fuzzy set classification of different levels of overdispersion of NBR n =5000 63

Table 4.11:Fuzzy set classification of different levels of overdispersion of CP n =20, 30 65

Table 4.12:Fuzzy set classification of different levels of overdispersion of CP n =50, 100 67

Table 4.13:Fuzzy set classification of different levels of overdispersion of CP n =200, 300 69

Table 4.14:Fuzzy set classification of different levels of overdispersion of CP n =500, 1000 70

Table 4.15:Fuzzy set classification of different levels of overdispersion of CP n= 5000 71

Table 4.16:Fuzzy set classification of different levels of overdispersion of GP
n =20, 30 73
Table 4.17:Fuzzy set classification of different levels of overdispersion of GP
n =50, 100 75
Table 4.18:Fuzzy set classification of different levels of overdispersion of GP
n =200, 300 78
Table 4.19: Fuzzy set classification of different levels of overdispersion of GP
n =500, 1000 80
Table 4.19: Fuzzy set classification of different levels of overdispersion of GP
n =5000 81
Table 4.20:Averaging Method for Determination of Threshold
n =20, 30, 50 85
Table 4. 21: Averaging Method for Determination of Threshold
n =100, 200, 300 86
Table 4. 22: Averaging Method for Determination of Threshold
n =500, 1000, 5000 90
Table 4. 23: Summary of the Min-Max Threshold values for the 4 count models 91
Table 4.24: The Threshold values for the model 91
Table 4. 25: Collinearity Statistics 91
Table 4.426:Fuzzy set method for classifying of difference in percentages of
Overdispersion 97
Table 4. 27: Parameter Estimation of Poisson Regression Model 97
Table 4. 28: Parameter Estimation of Negative Binomial Model 97
Table 4. 29. Parameter Estimation of Com-Poisson Model 98
Table 4. 30. Parameter Estimation of Generalised Poisson Model 98
Table 4.31. Akaike Information Criteria (AIC) 98

LIST OF FIGURES

Figure 4.1. Plot for number of Crashes	92
Figure 4.2. Plot for SPV	92
Figure 4.3. Plot for UPD	93
Figure 4.4. Plot for OVL	93
Figure 4.5. Plot for DGD	94
Figure 4.6. Plot for SOS	94
Figure 4.7. QQ-Plot for Poisson Model	99
Figure 4.8. QQ-Plot for NBR	99
Figure 4.9. QQ-Plot for CP	100
Figure 4.10 QQ-Plot for GP	100

ABBREVIATIONS

PO - POISSON

NB - NEGATIVE BINOMIAL

CP - COM-POISSON

GP - GENERALISED POISSON

VIF- VARIANCE INFLATION FACTOR

AIC- AKAIKE INFORMATION CRITERION

SPV- SPEED VIOLATION

UPD- USING PHONE WHILE DRIVING

OVL- OVERLOADING

DGD- DANGEROUS DRIVING

SOS- SLEEPING WHILE DRIVING.

CHAPTER ONE

INTRODUCTION

1.1 Background to the study

Count data are statistical data type in which the observations can only have non-negative integer values and the integers are generated by counting rather than ranking. They are made up of discrete numbers defined by a large number of elements or objects that are divided into various classes or groups. Such as the number of people infected with the Coronavirus (Cov-19), the number of people exposed to Ebola, the number of patients admitted to the hospital, the number of road transport accidents, the number of people involved in motor accidents and the number of children born by a woman. Count data achieved here because the number of the events involved with counting and the number of occurrences can be counted. Poisson distribution is commonly used to model count data distribution (different from normal distribution data) and it is commonly used for rare events. A count variable refers to the individual piece of count data. Poisson and Negative Binomial Distributions are the common distributions often used.

The Poisson regression model is a popular model for count data (Gschlobl 2013). It belongs to the class of Generalised Linear Model (GLM) because the error term is not normally distributed but Poisson distributed and also belongs to the Exponential class family. The mean of the model is assumed to be equal to its variance.

A popular assumption of Poisson distribution known as equi-dispersion is a phenomenon whereby the parameter of interest or location parameter is equal to its variance, that is, equality of mean and variance. Although, in real life settings, variance may not be equal to mean this leads to violation of one of the assumptions of Poisson distribution. The dispersion may be under-dispersion or over-dispersion. Whenever Poisson distribution is applied to count data and dispersion occurs, it is a problem. Why? This is because Poisson distribution has only one parameter and there is a need to model the extra-variation, that is why Consul and Famoye (1992), Famoye (1993), Winkerman and Zimmermann (1994) and Endo (2020), stated that Poisson distribution is too restrictive and inappropriate to handle dispersion, it is incapable of handling dispersion because it has only one parameter.

In count data, whenever variance is less than mean there will be under dispersion, this may be encountered in real life situations but under dispersion is not common as over-dispersion. Modeling under-dispersed data with a Poisson distribution will pose a problem and cause wrong estimation when under dispersed data are fitted with Poisson distribution. This makes the standard error to be over-estimated; which will lead to wrong conclusion about the predictors as not significant when indeed they are significant. Generalised Poisson among other models is commonly used in modeling under-dispersed data in the Poisson model. According to Consul (1989) "Generalized Poisson and Negative Binomial models are appropriate for modeling dispersion in count data".

Overdispersion underestimates the standard error, thereby leading to the wrong conclusion and in the effect making the parameters of predictors significant when they are not. It makes the standard error small when it is large; a misleading and wrong conclusion is given. Count data with many zeros are common in many applications (Bohning *et al* (1996), Rideout *et al* (1998), Agarwal (2002), Boanafede (2015). Apart from under dispersion and overdispersion, the problem of excessive zeros may occur in count data. That is, occurrence of too many zeros, Gardner *et al* (1995), German (2007), Guikema and Coffeit (2008), Lawless (2012), Gul Inan (2017), which may lead to Zero inflation. Whenever this occurs, standard error of the parameter is affected and the model will not be properly fitted. The model for excess zero consists of two generating processes; binary one that generates the structural zeros and the other which is the count from Poisson distribution (Lambert (1992), Lord *et al* (2005), Ver (2007), Lee (2012).

Many factors contribute over dispersion in count data; below are some itemised factors responsible for the overdispersion:

- Variability in the population of interest
- Omission of key or relevant predictors
- Incorrect functional form
- Outliers

The Poisson regression model is used to fit the data set and check if the violation of the assumption of the model exists, the ratio of residual deviance to its degree of freedom, if roughly equals one, is the evidence that Poisson regression model is adequate for the model, if not, then the alternative model is used Vuong (1989), Stram (1994), Stram (1995), Lindsey (1999), Marchini (2005), Richard (2007), Ismail (2007), Hinde (2008), Yang (2009), Ismail, *et al* (2013), Yan (2015). The question is, how can we detect overdispersion in count data? The rule of thumb is that when the ratio of residual deviance to its degrees of freedom roughly equals one, it means that dispersion does not exist and vice versa. Dispersion in count data is a problem to count data whenever it is modeled with Poisson regression model, The model will not be properly fitted, that is why, to account for the extra variation in count data, a model that can account for or allow for overdispersion is used as earlier stated by Winkerman (1994) and Nwankwo (2015). In the light of this, to account for overdispersion in count data, a model which has more than one parameter is used to take care of the extra variability.

One of the ways to handle dispersion is the introduction of additional parameter(s). Distributions that allow for dispersion are introduced. The commonly used distribution to address overdispersion is the Negative Binomial Distribution while Generalised Poisson Distribution handles both under and over dispersion. Nwankwo (2015) stated that “Negative Binomial regression addresses the issue of over-dispersion by including a dispersion parameter to accommodate the unobserved heterogeneity in the count data”. Negative Binomial distribution was developed by Green (1920), the Probability Distribution has two parameters which the dispersion parameter will cater for the extra variability in the count data Lord *et al* (2005), Famoye (2006), Greene (2008), Lee (2012). Apart from Negative Binomial Regression in handling the problem of overdispersion, Generalised Poisson regression also is used to address the problem of overdispersion. The distribution was developed by Nelder and Wedderburn (1972). It is a distribution with two parameters which is the dispersion parameter used to model the excess variability, Mccullargh (1989), Neeloon (2007).

Generalised Poisson distribution is one of the distributions used to model count data to address both the problem of underdispersion and overdispersion in count data Famoye *et al* (1993), Consul *et al* (1992), Ismail (2013). It is a generalisation of Poisson distribution. The model is applicable to handle the problem of overdispersion because of the extra parameter, unlike Poisson model with the single parameter. The model is a good competitor to Negative Binomial when modeling count data. The model has the advantage in that it can be used to model the count data without knowing the distribution of the count data.

Com-Poisson is a convolution of three distributions used to capture the problem of dispersion when it occurs with count data Shmuelli (2005), Kimberly (2010), Kimberly (2011). The distribution has an extra parameter, unlike Poisson distribution that has a single parameter. The extra parameter or dispersion parameter is used to capture variation or the excess variation but this distribution is not capable when excess zero is the cause of the problem. There was a dearth of research on this distribution Shmuelli (2005), Lord *et al* (2008, 2010).

When modeling excess zero in count data model, the Poisson regression and the common models that accommodate overdispersion may not be appropriate. A zero-inflated model will be applicable. Zero-inflation occurs when there are too many of zeros and in excess. For instance, to know the number of people that have been jailed in their lifetime, the majority of the respondents might have not been jailed before. In this case, a large number of zero which represent people who have not experienced such may be recorded. Even the study of the pandemic Corona Virus Disease (Covid-19) will have zero counts of the disease because a large number of Nigerians at the moment are not infected.

Whenever there is a scenario of too many zero counts, common distributions such as Negative Binomial and Generalised Poisson regression would not be adequate to model the count, in this case. It is needful to use a model such as Zero-Inflated Poisson (ZIP) model, Zero-Inflated Generalised Poisson model (ZIGP), Zero-Inflated Negative Binomial (ZINB) model, Zero-Truncated Model and Hurdle Models and a host of others to fit. If the conventional models are used they may lead to underestimation or overestimation of regression coefficients which result in invalid conclusions. Moreover, the standard error will be underestimated, and the null hypothesis will be rejected when it should have been accepted. When zeros occur in data sets, researchers or analysts deal with it as missing data, delete or impute it, but at times the zeros have meaning, importance, and should be considered as such Lindsey(1999).

Zero Inflated Poisson model is useful for modeling count data when there are excess zeros than expected; that is, when there are excess zero counts. It was developed by Lambert (1992) when he used the Poisson model for defects in manufacturing. It was discovered that the Poisson model provided a poor fit because of excess occurrence of zero count in the defect during manufacturing. Lambert (1992) stated that "Poisson regression is not appropriate for modeling count with excess zero`` and suggested that when modeling count data when there are excess zero, ZIP model will be better. Zero-inflated Poisson model is a mixture of both the zeros generating process, one for zero generation and the other governed the Poisson distribution which generates the count which may be zero. Apart from ZIP, Rideout *et al* (1996) suggested that other zero models could also be used because ZIP often provides a poor fit for handling excess zero. Likewise Famoye (1993) stated that "it is a motivation for developing ZIPG which is regarded as a better model than ZIP".

Zero-inflated Generalised Poisson (ZIPG) model is an alternative model to ZIP when there is evidence of overdispersion when dealing with excess zeros. The model is capable of handling count data when such problem arises. It is a rich family of ZIP and data generalisation of ZIP. The ZIP is sometimes incapable of capturing overdispersion and that gives room for modification of the ZIP for the ZIPG model. This model is capable of handling both underdispersion and overdispersion with count data when it is as a result of excess zeros. This model is good when the dependence of the count data is affected by some predictors. It is also an alternative model when ZINB fails and could not fit the data sets appropriately. According to Famoye (2006), "there are some cases in which ZIP may be inadequate and ZINB regression model may fail to converge".

Negative Binomial Regression model is an alternative model for handling overdispersion in count data, however, when there are too many zero counts than expected, negative binomial regression may not be adequate to handle the problem. Hence, there will be a need to use the zero inflated model for the data. Famoye (2006) stated that "overdispersion has the tendency to increase the proportion of zeros" and whenever there are too many zeros relative to the Poisson assumption, the negative binomial regression, and generalised Poisson regression tend to improve the fit of the data, Hinde (2008) also opined that ZINB regression is also good for modeling count data as an alternative.

ZINB is used to model count data where excess zeros and overdispersion occur. It is a combination of Negative Binomial and Logit distributions. The presence of excess zeros and the problem of overdispersion often occur with count data. Zero-Inflated Negative Binomial Regression was introduced to handle count data with excess zero because excess zero masquerades as over-dispersion (Greene (1994), Agarwal (2000) Hinde (1998). Due to the restrictive property of the Poisson model, several modifications and parameterisation have been done, developed, and suggested, the frequent among them is Negative Binomial regression. The model is only useful when there is a problem of overdispersion but not capable when there is excess zero in count data.

Zero-Truncated model consists of a response variable with non-zero. The value of the response variable can never be zero. The common scenario is the duration of stay by a patient in the hospital. The count starts immediately after admission, and hence, the response will not be zero. This model is useful when the count can never be zero which is suitable in real life settings.

Hurdle models for excess zeros are flexible and preferred to zero inflated models. They consist of two parts, one specifies the generating process for the counts in zero and the other process for positive counts. Both Hurdle and zero inflated models are used for modeling observation with the occurrence of zeros or excess zeros but the hurdle model is different in the way the data are analyzed and interpreted. A hurdle model, according to Cameron and Trivedi (2008) is a modified count model in which the two processes generating the zeros and positive are constrained to be the same.

Fuzzy set introduced by Zadeh L.A in 1965 is an extension of classic set. In a classic set, an element can either belong to a set or not but in fuzzy set, an element may belong or may not belong or partially belong to the set depending on the membership function. Fuzzy set introduced membership function to show the degree to which an element belongs to a set. The membership value lies in the interval of 0 and 1. Fuzzy set is useful when dealing with concept that has no sharp or precise boundaries. Fuzzy set has been applied to different fields of study; Bezdek (1981) applied fuzzy set into pattern recognition, Yang (1993) used fuzzy to propose and determine threshold for lower and upper bound of the parameter of two regressions, Javidi and Mansoury (2017) applied fuzzy set for selection and classification of gene into redundant and non-redundant.. In the view of this, fuzzy set will be relevant to this study, in order to classify overdispersion into different degrees of severity and determine the threshold of severity.

1.2 Statement of the Problem

Studies have shown that different models have been proposed and several parameter-

isations have been done to solve the problem of overdispersion associated with count, but discovered many a time that the problem of overdispersion is inherent in the count data. Count models are modified and alternative models are adopted at the slightest difference between mean and variance; even when the difference or the variation is negligible. This problem can be solved by re-structuring the model yet researchers still modify this model, but the question is: Is it a must to modify Poisson model when mean and variance are not equal? If yes, at what point should the modification be done? At what threshold is overdispersion severe? The threshold at which overdispersion is severe is ambiguous and fuizzy in nature. Overdisperson is a concept with unsharp boundaries , a piece of information that the boundary is not well clear-cut but this research would address this problem.

1.3 Justification

A lot of works on overdispersion have been done and different modified Poisson regression models have been developed without specifying the threshold for modification when overdispersion is severe. There is little or no contribution on when overdispersion is severe in the literature. Many modified the model without considering when taking the severity of overdispersion into account. In the research, Fuzzy set approach will be used to classify the different levels of severity of overdispersion.

Why Fuzzy Set Approach?

A fuzzy set theory is a powerful tool when classification is needed in a case where sharp criteria for determining membership function are not clear. The membership function is a key ingredient in a fuzzy set; it shows the degree of belongingness of each element to the universal set. Overdispersion is vague and appears blurry. "Fuzzy concepts" proposes that "somewhat vague terms" can be operated with, by assigning numbers to graduations of applicability.

According to Pawlus *et al* (2012), vagueness and lack of information can be successfully modeled by the Fuzzy method. Fuzzy set theory is used because it models specific types of uncertainty under specific types of circumstances (Zimmermann 2001), hence it can be used in categorising the different degrees of overdispersion. The basis of every fuzzy model is the membership function (Bezdek, 2014).

Fuzzy sets theory has been extensively used to relax or generalise classical methods from a dichotomous to a gradual characteristic fuzzy set, the characteristic function allows various degrees of membership for the elements of a given set (Zimmermann 2001).

1.4 Motivation for the Study

Many researchers have proposed different models to handle the problem of dispersion in count data. However, of these models, there is still a dearth of information about when dispersion is a serious cause of concern. Some modified at 10% while some modified at 75% Lambert (1992). The point and percentage at which researchers modify their models are of interest in this study. Is it necessary to always modify the models when there is dispersion? If yes, at what level or percentage is the modification necessary?

1.5 Aim and Objectives

This study aims to determine a threshold for modification of some count models (Poisson, Negative Binomial, Conway, and Generalised Poisson models) when the problem of overdispersion is inherent.

The specific objectives are:

i. To develop a method using fuzzy set to categorise the different levels of overdispersion.
ii. To determine at what point is there need to modify a model (Poisson, Negative Binomial Conway, and Generalised Poisson models) when the problem of overdispersion is inherent.
iii. To investigate the effect of the identified threshold on life study.

1.6 Significance of the Study

This research will contribute to knowledge when there is a problem of overdispersion and at what point and threshold it is needful to modify Poisson, Negative Binomial, ComPoisson, and Generalised Model when considering count data. The research made use of secondary data collected or extracted from the yearly publication of Federal Road Safety Corps Nigeria, from the year 2014 to 2018. Five variables were used to model the count to know if those variables contributed to the study and their usefulness for further studies.

1.7 Organisation of the Dissertation Presented

This dissertation consists of five chapters. Chapter one discusses the general introduction of count data, the problem of overdispersion associated with count data, different models for modeling the count data, motivation of the study, problem statement, objectives of the study, and the significance of the study.

Chapter two is the literature review on count data and overdispersion models. Theoretical Framework of the models with the conceptual discourse of the study with application to life study.

Chapter three deals with the research methodology and design. Chapter four presents the data and analysis with its interpretation of the results. Chapter five discusses the summary of the findings, conclusions and makes recommendations, and suggests further studies.

CHAPTER TWO

LITERATURE REVIEW

2.0 Chapter Overview

In many real-life settings, count data exhibit overdispersion. Poisson regression is a choice model for count data that can only handle problems with a single parameter. Poisson regression model, though a common model with count data, fails to model the data appropriately because it is restricted, inappropriate, and limited in handling overdispersion in count data. It is the simplest model but many a time not adequate to model the data. A Poisson distribution has a single parameter which is the parameter of the distribution which is the mean, there is a commonly known assumption that mean equals the variance i.e equidispersion.

This distribution is limited, restricted, and inappropriate to capture overdispersion with the count data Winkermann (1994), Greene (1994), Famoye (2004), Nwankwo (2015), Lambert (1992), Kimberly *et al* (2010). The presence of overdispersion in count data poses a problem in count data because it will lead to a poor fit, invalid, and incorrect conclusion of the result. It is important to discuss briefly here different distributions that have been used to model count data to capture excess variation. First, the simplest model will be discussed and the alternative models which are better to accommodate over-dispersion will later be discussed. One of the usual assumptions of Poisson distribution known as equi-dispersion which is equality of mean and its variance; count data often manifest overdispersion, when this occurs the variance is greater than the mean. When overdispersion occurs in the that data, Poisson model, Poisson regression will be inappropriate for modeling count data but appropriate only when there is no problem of dispersion. When the data do not exhibit dispersion, Poisson model can be used to fit the data. The presence of dispersion in count data results in a lack of fit in Poisson regression.

The Poisson regression model is a common model for count data (Gsclobi 2013). Poisson regression is a simple method to model count data but only when the assumption still holds. This model will not be adequate to model the count data when overdispersion exists with the count data (Winkermman (2003), Famoye (1993), Famoye (2004), Famoye (2006), Greene (1994), Nwankwo (2005), Greene (2008) Wang *et al* (1997). The single parameter of Poisson is limited and not capable when overdispersion and excess zero are inherent in count data, it will fail to model the extra-variation in case it appears. To properly model the count data with overdispersion, an extra parameter is introduced to the model. Different models have been used to capture the problem of overdispresion and several parameterizations have been done to solve the problem of overdispersion and excess zero.

Poisson regression has been applied to different fields of study such as transportation, insurance to study the number of claims by the clients and the severity of the claims, number of admitted and discharged patients, and number of students admitted to a higher institution to mention but a few (Kimberly *et al* (2010, 2011), Ozmen (2007), Rensham (1994).

With real-life data, using the Poisson regression model, there is usually evidence of overdispersion in count data. The single parameter of Poisson regression makes it inappropriate to handle the extra variation in the model. Winkelmann (1994, 2008) stated that "Poisson distribution is too restrictive and inappropriate to handle dispersion, it is incapable to handle dispersion because it has only one parameter". Likewise, Nwankwo (2015) also stated that "Negative Binomial regression addresses the issue of overdispersion by including a dispersion parameter to accommodate the unobserved heterogeneity in the count data".

Apart from Negative Binomial Distribution, other catalogs of distributions have been developed to tackle or solve the problem of overdispersion in count data. According to Consul (1989) "Generalized Poisson and Negative Binomial are appropriate for modeling dispersion in count data". Generalized Poisson is also appropriate to fit the model of count data when over-dispersion arises. The usefulness of Generalized Poisson is that it can be used to fit the model even when the probability function of the count variable, that the response variable (Y), is unknown, as far as there is equality of the Poisson mean and variance. The occurrence of over-dispersion in count data leads to underestimation of the standard errors and overstates the significance of Poisson regression parameters.

There are other mixtures of Poisson regression models that have been suggested to address the problem of over-dispersion in count data. The model served as alternative to handle dispersion in case it arises. Such models include Com-Poisson, Poisson-Lindley distribution, Poisson-inverse Gaussian (PIG), Poisson-lognormal (PLN) Srinivas (2008). Many researchers have proposed different models to overcome the problem of overdispersion and application of the model to real life situations to show the efficiency and importance of the proposed, developed, and modified models.

The approach for handling overdispersion is different when the problem of excess zero is encountered. At this point, the Zero-inflated model is opted for instead of the conventional Poisson model. An extensive discussion of what different researchers have done to address the problem of overdispersion in count data will be considered in the next section.

2.1 Terminologies

Dispersion: This is the greater variability presented in the data sets than expected in the given model. The dispersion may be underdispersion or overdispersion.

Equidispersion: It is the equality of the mean and variance. It is one of the popular assumptions of Poisson distribution.

Underdispersion: This phenomenon occurs when variance is less than mean.

Overdispersion: It occurs when the variance is greater than the mean.

Dispersion parameter: It is the extra or additional parameter introduced in the distribution to capture the extra -variation in count data when the Poisson model fails to properly fit the model.

Excess zero: It is the presence of too many occurrences of zero than expected in count data.

Hurdle model: It is a model which consists of two parts, one account for zero counts and the other accounts for the distribution of the non-zero.

Link function: It is a function that allows the incorporation of explanatory variables in a model.

Threshold: it is a value that serves as a benchmark or yardstick.

A fuzzy set is a set that consists of elements that can belong, partially belong or totally belong to the set.

Membership function: It is a function that shows the degrees of closeness of elements to the sets. It is defined as $\mu_s(D) \rightarrow [0.1]$ where the element of set D is mapped to value between 0 and 1.

2.2 Theoretical Framework

Poisson Regression Model

The probability density function is given as

$$P(Y = y) = \frac{e^{-\theta}\ \theta^{y}}{y!}, \quad y = 0,1,2,3..., \qquad \theta > 0 \tag{2.1}$$

Mean = $E(X/Y) = \theta$ (2.2)

Variance = $Var(X/Y) = \theta$ (2.3)

$$E(X/Y) = Var(X/Y) = \theta \tag{2.4}$$

Where y_i is the count variable and θ is the parameter of interest which is also equal to its variance(σ^2).

$$\log(\theta) = logE(Y/X) = \alpha + X'\beta \tag{2.5}$$

where $X'S$ are independent variables, $\alpha\ \varepsilon\ R\ and\ \beta\ \varepsilon\ R$. The mean is modeled as log-linear function of the observed variables, that is, $log(\theta) = logE(Y/X) = \alpha + X'\beta$ and the link function, $\theta = exp(X'\beta)$. Poisson regression belongs to the class of exponential family and generalised linear model.

Poisson Model Assumptions

The following are the underlying assumptions of Poisson regression model;

(1) The response variable is non-zero integer.

(2) It has a single parameter.

(3) The parameter of interest (μ) is the mean.

(4) Equi-dispersion; mean equal its variance.

(5) The observations are independent

2.3 Parameter Estimation of Poisson Regression Model

In Statistics, statistical inference is carried out to estimate the parameter of the population because the characteristic of the population parameter is constant and unknown. The estimate of sample would be used to infer to the characteristics of the population.

2.4 Maximum Likelihood Estimation of Poisson Regression Model

In the Poisson model, many methods are used for the estimation of the parameter of the model but in this study, the maximum likelihood will only be discussed. It is the method of estimation used in Statistics for estimating a parameter of interest. It is done by maximizing the likelihood function of the parameter of the model. It is easy to obtain the log-likelihood function given as

$$l(\beta; y) = \ln L_n(\beta; y) \tag{2.6}$$

Let $Y_1 .., Y_n$ be independent random variables from a population with the probability density function $f(y/\theta_1,...\theta_n)$, the likelihood function is defined by

$$L(\theta / y) = L(\theta_1, ...\theta_n / y_1,...y_n) \tag{2.7}$$

$$= \prod_{=1}^{n} f(y/\theta_1, ..., \theta_n), \quad \text{where yi is the sample} \tag{2.8}$$

Let $Y_i, i = 1, .. n_1$, be independent random variables distributed as Poisson with the Probability density function as given in (2.1);

$$P(Y = y) = \frac{e^{-\theta} \theta^{y}}{y!}, \quad y = 0, 1, 2, 3..., \quad \theta > 0$$

The log- likelihood is obtained by

$$L(\theta_1, ...\theta_n) = \prod_{i=1}^{n} \frac{e^{-\theta_i} \theta_i^{y_i}}{y_i!} \tag{2.9}$$

The parameter θ is defined in terms of $\theta_1,...\theta_n$ and the covariates $x_{i1},\ldots x_{in}$ via the link function. Taking x_{0i} =1, the log-likelihood is therefore given as

$$L(\theta_1,...\theta_n) = \sum_{i=1}^{n} y_i \log\theta_i - \theta_i - \log y_i! \tag{2.10}$$

$$= \sum_{i=1w}^{n} y_i \left(\sum_{i=0}^{n} \theta_i x_m\right) - e^{\left(\sum_{i=0}^{n}\theta_i x_m\right)} - \log y_i! \tag{2.11}$$

The maximum likelihood estimator solution by taking m = 0,…k,

$$\frac{\partial l}{\partial \theta_m} = \sum_{i=1}^{n} x_{im}\left(y_i - \ell^{\sum_{u=1}^{p}\theta_u x_{iu}}\right) \tag{2.12}$$

Solution to (2.12) can be obtained numerically by the Newton- Raphson

2.5 Com-Poisson Regression Model

Conway et al (1962) developed the model called Com-Poisson which is also known as Conway which was named after Conway. The Com-Poisson is a generalization of Poisson distribution by the introduction of extra or additional parameters into the Poisson model to capture both under and overdispersion in count data in case the Poisson model is not adequate. It belongs to the exponential family. It is the convolution of three distributions namely- Poisson, geometric, and Bernoulli distributions. Geometric distribution as a special case and Bernoulli distribution as a limit case. This distribution was used to proffer a solution to the problem of the queuing system when Poisson distribution could not tackle the problem as a result of the extra variation.

2.6 Probability Density Function

Given a Com-Poisson distribution the Probability mass function is given as

$$P(X = x) = f(x,\lambda,v) = \frac{\lambda^v}{(x!)Z(\lambda,v)} \quad \text{where} \quad Z(\lambda,v) = \sum_{J=0}^{\infty} \frac{\lambda^j}{(j!)^v} \tag{2.13}$$

X is the random variable, the function $Z(\lambda, v)$ serves as a normalization constant. Hence the

$$Mean = \sum_{j=0}^{\infty} \frac{j\lambda^{j}}{(j!)Z(\lambda, v)} \text{ and } Variance = \sum_{j=0}^{\infty} \frac{j^{2}\lambda^{j}}{(j!)Z(\lambda, v)} \qquad (2.14)$$

The parameter v is the additional parameter while λ is the parameter of the Poisson distribution. It is discovered here that mean and variance are not equal. In order to modify Poisson model, another parameter was introduced into Poisson distribution through convolution of the aforementioned distributions, so that the extra parameter will model the extra-variation. And when the parameter (v) equal 1, the Com-Poisson becomes Poisson distribution, when v tends to infinity it becomes a Bernoulli distribution, and Geometric distribution when $v = 0$. There was no or little research on the distribution until when it was revived by Shmueli in 2005 (Shmueli 2005). Recently, many researchers have explored the distribution to model count data, used to capture under and overdispersion in count data, yet, the model has not fully been explored and the distribution is limited when excess zero is encountered and the threshold for the need for modification was not specified.

2.7 Maximum Likelihood Estimation of Com-Poisson Regression Model

Likewise, the maximum estimation can also be done for Com-Poisson regression as it is done for the Poisson model when considering the estimation of the parameter of the model. Let $Y_i, i = 1, \dots n,$ be independent random variables distributed as Com-Poisson regression model with the Probability density function given below

$$P(X = x) = f(x, \lambda, v) = \frac{\lambda^{v}}{(x!)Z(\lambda, v)}$$

The log- likelihood for the observation (Y= y_i =1,..n_i) can be written as

$$\log L = \sum_{i=1}^{n} y_i \log \lambda_i - v\sum \log y_i - \sum \log Z(\lambda_i, v)_i \qquad (2.15)$$

Under the constraint of v ≥ 0 the estimate of the maximum likelihood can be obtained by directly maximizing (2.14).

2.8 Negative Binomial Regression Model

Different modifications have been done for the Poisson model to accommodate overdispersion in count data. The Negative Binomial Regression model is commonly used to model count data in case count data suggests overdispersion (Nwankwo (2005), Consul *et al* (1989), Green (1998), Lord *et al* (2005). Negative Binomial Distribution was developed by Greenwood (1920). This model is used to address overdispersion because it is a distribution with extra parameters, unlike Poisson model which has a single parameter. Here, the modification from Poisson to Negative Binomial Model is considered.

Let Y be a random variable distributed as Poisson distribution with the Probability Density Function

$$P(Y=y)=\frac{e^{-\lambda}\lambda^{y_i}}{y!},\ \mathrm{y}=0,1,2,3....,\qquad \lambda>0 \tag{2.16}$$

$$Mean = E(X\ /Y)=\lambda \tag{2.17}$$

$$Variance = Var(X\ /Y)=\lambda \tag{2.18}$$

$$E(X\ /Y)=Var(X\ /Y)=\lambda \tag{2.19}$$

where y_i is the count and λ is the parameter of interest which is also equal to its variance for incorporation of the covariates given by $\log(\lambda)=\log E(Y/X)=\alpha+\beta' X$ with independent variables, $\alpha \varepsilon R$ and $\beta \varepsilon R$. The mean is modeled as log-linear function of the observed variables i.e $\log(\lambda)=\log E(Y/X)=\alpha+\beta' X$ and the link function, $\lambda=\exp(X'\beta)$. Unobserved latent heterogeneity can be incorporated into the conditional mean and restructure the functional form such as

$$E(Y/\theta)=e\,(x_i\beta+\theta_i)=w_i\lambda_i \text{ given } w_i=e^{(\theta)} \tag{2.20}$$

follows Gamma distribution with mean 1 and variance $V^{-1} = c$ with Probability Density Function

$$f(w_i) = \frac{V^v w_i^{v-1} e^{(-vwi)}}{\Gamma(v)}, \quad w_i \geq 0, \quad V. > 0 \tag{2.21}$$

The conditional Poisson regression is

$$P(Y_i = y_i) = \frac{e^{(-vw_i)}(w_i \lambda_i)^{y_i}}{y_i\,!}, \qquad y_i = 0,1,\ldots \tag{2.22}$$

Thus, the Negative Binomial regression with Probability Density Function is given as

$$P(Y_i = y_i) = \frac{\int e^{-(-vw_i)}(w_i\lambda_i)^{y_i} Vve^{-(-vw_i)}(w_i\lambda_i)^{y_i}}{y_i\,!} \tag{2.23}$$

$$= \frac{\Gamma(y_1 + v)}{\Gamma(yi)\,\Gamma(v)} \left(\frac{v}{v+\lambda_i}\right)^{y_i} \left(\frac{\lambda_i}{v+\lambda_i}\right) \tag{2.24}$$

By modifying equation (2.16), replace λ_i by $w_i \lambda_i$, it is seen from (2.20) that an additional parameter which is ω_i has been introduced into Poisson model in order to modify the model to have a model which will accommodate overdispersion. Where

The mean is

$$E(Y_i) = \lambda_i \tag{2.25}$$

and the variance is

$$Var(Y_i) = \lambda_i(1+\theta^{-1}) = \lambda_i(1+c\lambda_i) \tag{2.25}$$

and

$$\theta^{-1} = c \tag{2.26}$$

c denotes the dispersion parameter equation (2.22) referred to Negative binomial regression (2) by Cameron (1986). Both Negative Binomial Regression 1 and 2 was done by reparamazation of Poisson model, in-order to produce NB-1 another parameterizationcan be done by letting

$$\theta = c^{-1}\lambda \qquad (2.27)$$

to have Negative Binomial Regression 1 with

$$E(Y_i) = \lambda_i \qquad (2.28)$$

and the variance is

$$Var(Y_i) = \lambda_i (1 + c) \qquad (2.29)$$

for carrying out statistical test for Negative Binomial regression sometimes it is difficult to choose a better model of the Negative binomial regression. Mostly, scholars encountered selection problem when using Negative binomial regression model to fit count data. This arises because the problem of selection between Negative binomial regression model 1 and Negative binomial regression model 2. The functional form of Negative binomial regression model is done in which it is nested was introduced by Greene (2008) such as selection can easily be done when carrying out the statistical test. This functional form developed by Greene (2008) known as Negative binomial regression model – $P(NB - P)$, by letting

$$\theta = c^{-1}\lambda_l^{2-p} \qquad (2.30)$$

The men is

$$E(Y_i) = \lambda_i \qquad (2.31)$$

and the variance is

$$Var(Y_i) = \lambda_i (1 + c\lambda_i^{p-1}) \qquad (2.32)$$

where c is the dispersion parameter and P is the functional parameter, when $P = 1$ it is Negative binomial regression -2 and when $P = 2$ it is Negative binomial regression-2. Maximum likelihood method is employed for the parameter estimation.

2.9 Maximum Likelihood Estimation of Negative Binomial Regression Model

Let $Y_i, i = 1, \ldots n,$ be independent random variables distributed as Negative Binomial Regression Model with the Probability density function given in (2.22)

$$\frac{\Gamma(y_1 + v)}{\Gamma(y_i)\,\Gamma(v)} \left(\frac{v}{v+\lambda_i}\right)^{y_i} \left(\frac{\lambda_i}{v+\lambda_i}\right)$$

The log- likelihood for the observation (Y_i = y_i = 1,..n_i) can be written as

$$L(\beta, y, v) = \sum_{i=1}^{n} y_i \log\left(\frac{ve^{(x\ \beta)}}{1-ve^{x'\beta}}\right) - \frac{1}{v}\log(1+ve^{x'\beta}) + \log\Gamma(y_i) + \frac{1}{v}) - (\log(y_i + 1) - \log(\frac{1}{v})$$

(2.33)

the maximum likelihood estimates $(\hat{\beta}, \lambda)$ may be obtained by maximazing $l(\hat{\beta}, \lambda)$ with respect to β and λ Thus;

$$\frac{\partial l(\beta,\lambda)}{\partial \beta} = \frac{\sum_i (y_i - \lambda)x_i}{1 + v\lambda_i} = 0, \qquad j=1,2\ldots p \qquad (2.33)$$

and

$$\frac{\partial l(\beta,\lambda)}{\partial \lambda} = \sum_i \left\{ \sum_{i=1}^{y_i - 1} \left(\frac{1}{1+\lambda v} + v^{-2} \log(1+v\lambda \right) - \frac{(y_i + v^{-1})\lambda_i}{1+v\lambda_i} = 0, \qquad (2.34)$$

2.10 Generalised Poisson Regression Model

Apart from Negative binomial regression model for handling overdispersion in count data, Generalised Poisson Model is also used to model count data for both underdispersion and overdispersion. Oftentimes, Negative Binomial Regression model may not converge when used to fit count data and the source is not as a result of excess zero. According to Consul (1989)``the commonly used distribution to address overdispersion is the Negative Binomial Distribution while Generalized Poisson Distribution handles both under and overdispersion.``

The Probability mass function of Generalised Poisson Distribution is given as

$$\Pr(Y = y_i) = \frac{\mu(\mu + \alpha y_i)^{y_i - 1} e^{(-\mu_i - \alpha y_i)}}{y_i!} \qquad (2.35)$$

where $\mu > 0$ and max(-1, $\frac{-\mu}{4}$)< $\alpha < 1$.

The mean $E(Yi) = \theta = (1-\alpha)^{-1}\mu$ and the variance Var(Y_i)=(1-α)$^{-3}$$\mu = (1-\alpha)^{-2}$ where $(1-\alpha)^{-2}$ denotes the dispersion factor and α dispersion parameter. This distribution is often used to model count data whenever there is problem of dispersion whether it is underdispersion or overdispersion. Different parameterizations also have been done for Generalised Poisson Regression model namely Generalised Poisson Regression model 1(GP1) and Generalised Poisson Regression model 2 (GP2), Generalised Poisson Regression model 1 (GP1) can be developed by letting $\mu = (1-\alpha)\lambda_i$ in (2.34), Generalised Poisson regression with mean $E(Y_i) = \lambda$ *and variance* $Var(Y_i) = (1-\alpha)^{-2}\lambda_i$ When $\alpha = 0$ the model reduces to Poisson regression model when it is $\alpha \prec 0$ the model can accommodate underdispersion while $\alpha \succ 0$ the model can accommodate overdispersion.

Another form of parameterization is also done for GP1 by letting $\alpha = c(1+c)^{-1}$ and $\lambda_i = (1=c)^{-1}\lambda_i$ in (2.34) with mean $E(Y_i) = \lambda_i$ and variance $Var(Y_i) = (1+c)^2\lambda_i$ while c is the dispersion parameter. Generalised Poisson regression model 2(GP2) is produced by doing another parameterization by $\alpha = (1+\lambda_i c)^{-1} c\lambda_i$ and $\mu = (1+\lambda_i c)^{-1}\lambda_i$ in (2.34) with mean $E(Y_i) = \lambda_i$ and variance $Var(Y_i) = (1+c\lambda_i)^2\lambda_i$. Both Generalised Poisson regression model 1(GP1) and Generalised Poisson regression model 2 (GP2) are not nested and selection between the two models is difficult like the case of Negative binomial regression model 1 and 2.

Another parameterization is also done which is known as Generalised Poisson regression model-P (GP-P) by letting $\mu = (1+c\lambda_i^{p-1})\lambda$ and $\alpha = (1+c\mu_i^{p-1})^{-1}c\lambda_i^{p-1}$ in equation (2.34), a functional form of Generalised Poisson regression model is produced known as Generalised Poisson regression model-P (GP-P) with the mean $E(Y_i) = \lambda_i$ and variance $Var(Y_i) = (1+c\lambda_i^{p-1})^2\lambda_i$ where c is the dispersion parameter and P is the functional form. When $c = 0$ Generalised Poisson regression model-

P(GP-P) is reduced to Poisson regression model , when $c < 0$ the model allows for underdispersion and overdispersion when $c > 0$

Moreover, when P=1. Generalised Poisson regression model-P(GP-P) is reduced to Generalised Poisson regression model 1(GP1), and Generalised Poisson regression model 1(GP2) when P=2. Both Generalised Poisson regression model GP1 and GP2 are not nested but GP-P parametrically nests both model so that statistical test can be done and tested against the alternative model. These models are limited and cannot model the count data when there is problem of excess zero. There is a need to consider those models which can accommodate excess zero with count data.

2.11 Maximum Likelihood Estimation of Generalised Poisson Regression Model

Let consider Y_i, i =1,..n_1, be independent random variables distributed as Generalised Poisson Regression Model with the Probability density function given in (2.34) ;

$$\Pr(Y_i = y_i) = \frac{\mu(\mu + \alpha y_i)^{y_i - 1} e^{(-\mu_i - \alpha y_i)}}{y_i!}$$

The log-likelihood for the observation (Y=y_i=1,…n_i) can be written as

$$L(\alpha, \beta) = \prod_{I-1}^{N} \frac{\mu_i(\mu_i + \alpha y_i)^{y-1} e^{(-\mu_i - \alpha y_i)}}{y_i!} \tag{2.36}$$

$$L(\alpha, \beta) = \sum_{i=1}^{n} y_i(\ln e^{xi\beta}) - (y_i - 1)\ln(1 + \alpha y_i) - \ln y_i! - e^{(-\mu_i - \alpha y_i)} \tag{2.37}$$

2.12 Zero-Inflated Model (ZI Model)

They are many different types of zero inflated models because different models can handle different types of overdispersion as a result of excess zero. The models include Zero-inflated Generalised Poisson, Zero-inflated Negative Binomial, Zero- Truncated Model.

2.13 Zero-Inflated Poisson Model (ZIP MODEL)

Zero-inflated Poisson was developed by Lambert (1992) when modeling for the defect in manufacturing. It was discovered that the Poisson model provided a poor fit because of the excess occurrence of zero count in the data used to model defects during manufacturing, when Modeling count data with too many occurrences of zero, zero-inflated Poisson model will be better. Apart from ZIP, Rideout *et al* (1995) suggested that other models could be used because ZIP often provides a poor fit for handling excess zero, when it was used to fit the data there was evidence of overdispersion with the count data, this motivated Famoye (2006) to develop Zero-Inflated Generalised Poisson model ZIP is obtained by mixing a distribution degenerated at zero with the Poisson distribution which accommodates explanatory variables in both the zero process and the Poisson distribution. The Probability mass function of Zero-inflated Poisson model regression is

$$\Pr(Y_i = y_i = \begin{cases} \theta_i + (1+\theta_i)e^{-\lambda_i}, & y_i = 0 \\ (1-\theta_i)\dfrac{\lambda_i}{y_i} e^{-\lambda_i} & y_i > 0 \end{cases} \tag{2.38}$$

Where $0 \le \theta_i \le 1$ and $\lambda_i > 0$ mean given as $E(Y_i) = (1-\theta_i)\ \lambda_i$ and Variance, $Var(Y_i) = (1-\theta_i)\lambda_i((1+\theta_i)\ \lambda_i$. Zero-inflated Poisson model reduces to Poisson regression model when $\theta_i = 0$. To incorporate covariate both log link and logit link can be used as the link function.

2.14 Zero-Inflated Generalised Poisson Model (ZIPG)

It is an alternative model to ZIP when there is still evidence of overdispersion while dealing with excess zeros. The model is capable of handling count data when the problem arises. It is a rich family of ZI models. It is a generalization of ZIP; ZIP is sometimes incapable of capturing overdispersion that gives room for modification of the ZIP for ZIPG. This model is capable of handling both underdispersion and overdispersion with count data when it is a result of excess zero.

This model is good when the dependence of the count data is affected by some predictors. It is also an alternative and good competitor to Zero-Inflated Negative Binomial Model, whenever ZINB could not fit the data appropriately. According to Famoye (2006)," some cases where ZIP model was not inadequate and Zero-Inflated Negative Binomial Model may fail to converge as a result ZIPG will be suitable for this purpose". The Pdf of ZIGP Is given as

$$\Pr(Y = y_i / x_i ;.\alpha) = \begin{cases} \theta_i + (1+\theta_i)\ f(\mu,\alpha,y_i), & y_i = 0 \\ (1-\theta_i) f(\mu,\alpha,y_i) & y_i > 0 \end{cases} \qquad (2.39)$$

where y_i = 0,1,2,…. and $f(\mu,\alpha,y_i)$ are the Pdf of Generalised Poisson Distribution.

2.15 Zero-Inflated Negative Binomial Model (ZINB)

Negative Binomial Regression is an alternative model for handling overdispersion in count data however, whenever there are too many zeros than expected in count data Negative Binomial Regression may not be adequate to handle the problem. ZINB was introduced to handle count data with excess zero because excess zero masquerades as over-dispersion Greene (1994).

Several modifications have been proposed due to the restrictive property of the Poisson model, the most frequent among them is ZINB. ZIP is useful in modeling count data when there are excess zeros than expected i.e when there are excess zero counts but sometimes it may not fit the count data with excess zero very well. ZINB is used to model count data where excess zeros and overdispersion occur. It is a combination of Negative Binomial and Logit distribution. The presence of excess zeros and the problem of overdispersion often occur with count data. The Pdf of ZIGP Is given as

$$\Pr(Y = y_i) = \begin{cases} P + (1-P_i)(1+\vartheta/r)^{-r}, & y_i = 0 \\ (1-P_i)\dfrac{\Gamma(y-1)_i}{y!\Gamma(r)}(1+\vartheta/r)^{-r}(1+r/\vartheta)^{-y_i} & y_i > 0 \end{cases}$$

(2.40)

where P Is the proportion of structure zero, is the probability of success and r is the dispersion parameter

2.16 Zero-Truncated Model

Zero-Truncated model consists of the response variable with non-zero. The value of the response variable can never be zero. The common scenario is the duration of stay by a patient in the hospital the count started immediately after admission here the response will not be zero. This model is useful when the count can never be zero. The Pdf of ZIGP Is given as

$$\Pr(Y = y_i) = \frac{\Pr(Y = y)}{\Pr(Y > 0)}, \quad \text{y=1,2,3,...} \qquad (2.41)$$

2.17 Hurdle Models for Excess Zeros

Hurdle models for modeling zero count are flexible and preferred to zero-inflated models. They consist of two parts; one specifies the generating process for the counts in zero and the other specifies process for positive counts. Zero-inflated and hurdle models of count data with extra zeros for example are used when there is a problem of excess zero. The hurdle model is a modified model for observation with occurrences of excess zero count, both hurdle models and zero-inflated models are used for this purpose but hurdle commonly used when there are excess sampling zeros.

$$\Pr(Y = y_i / x_i;.z_i) = \begin{cases} f_1(y_i / z_{i\,i}), \beta_{1_i} & y_i = 0 \\ \phi\ f_2(y_i / x_{i\,i}), \beta_{2\,i} & y_i > 0 \end{cases} \qquad (2.42)$$

where the numerator ϕ of represent the probability of crossing the hurdle.

2.18 Literature Review on Count Data Models

Winkelmann and Zimmermann (1994) noted that the restriction imposed on the Poisson model because of the assumption of the mean-variance relationship limited the Poisson model. The work of King (1989) was extended by the introduction of extra parameter k into the model to have a wide class of generalized event count (GEC^k) which accommodates both over and underdispersion. This model was applied to German data on fertility, divorce, and mobility. The new model was equally applied to the data sets used by Winkelmann (2004*) et al.* Therefore, the use of General GEC^K was suggested instead of Poisson regression to model count data. Model is limited in sense that it failed to consider overdispersion when is of serious concern.

Lambert (1992) Extended the work of Heibron (1989); noted Poisson model and the existing models failed to model count data with too much of zero, therefore, proposed Zero-Inflated Poisson (ZIP) model to account for excess zero counts. In the work, ZIP regression with an application to defects in manufacture fitted the data sets with Poisson model that there was evidence of excess zero in the data, 75% of ‘zero’ were recorded among the data sets, the ZIP according to the Lambert (1992) could accommodate zero counts and preferred to Poisson, negative binomial, Generalised Poisson e.t.c.

Consul and Famoye (1992) modified the Generalised Poisson model developed by Consul *et al* (1973) to model the count data because the model still fails to accommodate some levels of overdispersion, especially when the case of excess zero comes in. However, many researchers suggested that the model is good for modeling both under and overdispersion without knowing the probability function of the response, it has the advantage of fitting the model without knowing the probability function of the response, as far as the mean is equal to the variance of the Poisson model.

Greene (1994) proposed Zero inflated Negative Binomial Regression (ZINB). Several modifications of both the Poisson model and Negative binomial distribution for modeling the count data were presented. The problem of overdispersion can be solved when Negative binomial regression is used. ZINB was proposed when the existing models could not handle the problem of too many zero counts, out of 1319 sample, 1060 zero occurred in data sets of number of consumers that reported the credit card agency, 10 The new model was applied to study consumer credit behaviour. Test to differentiate between overdispersion and excess zero was also presented. The maximum likelihood method was used for parameters estimation, Vuong's statistic was used for the statistical inference.

Cameron and Trivedi (1996) pointed out that the assumption of equidispersion may not hold most of the times in real life settings. When the real life data are used, the assumption may fail. The research presented different functional forms of the Negative Binomial regression model. The difference between the forms of Negative binomial regression was considered but did not suggest the parametric test to choose between the two variants of the model. Two different functional forms of Negative binomial were developed which were labeled Negative Binomial one (NB1) and Negative Binomial two (NB2). The model was applied to health insurance consultation to doctor between the periods of two weeks, to know if health insurance lncrease the frequency of utilizing the heatlh . It was discovered the data sets were overdispersed; the sample mean was 0.302 and the sample variance was 0.637. The conducted tests revealed that both models were better than the Poisson model but could not provide a viable statistical test to select between the kinds of the Negative binomial.

Rideout *et al* (1998) discussed and examined various models that can accommodate excess zeros in count data. Different count models were used to examine Horticultural study. The conventional models like Poisson, Negative Binomial, Generalised Poisson usually fail to account for excess zero in count data and commented that it is better to use Zero-Inflated models to handle the problem. In the study, it was discovered that both Poisson and Zero-inflated Poisson (ZIP) provided a poor fit, and eventually modeled the data sets with Zero-Inflated Negative Binomial (ZINB). ZINB was found better to account for the problem of excess zeros. The maximum likelihood method was used for parameter estimation and the score test was used for the Statistical inference. ZINB was suggested for the problem of excess zero in count data, it is better to model the count with ZINB, which was also supported by Lambert (1992) but Famoye (2006) said that many a time ZINB fail to capture this problem of excess zero as a result proposed a competitive model to ZINB.

Lindsey (1999) presented a contrary view of detecting overdispersion in count data. Assumptions of overdispersion were made about the count data without considering the evidence of present overdispersion in the count data. In studying how should overdispersion be modeled? The rule of thumb is that the ratio of the residual deviance to the degree of freedom should be equal to one for the presence of overdispersion in the count data. The study suggested and proposed that deviance should be twice the degree of freedom should be considered, because sometimes Poisson model may model the data set appropriately. However, no suggestion or recommendation is made on when overdispersion is severed or not.

Famoye (2004) applied Generalised Poisson regression (GPR) model to life study of accident data of Alabama department of Public safety records. The model was used to assess the relationship between the occurrence of road crashes and some predictors. There was a case of overdispersion when the sample mean and sample variance were examined; the estimate was 0.76 and 1.33 respectively. GPR performed better than the Poisson model and it is a good competitor to other models that accommodate overdispersion in count data. Deviance statistic was used to account for the presence of overdispersion; Akaike information criterion was used to measure the adequacy of the model. GPR reduces to the Poisson model when the dispersion parameter converges to zero. However, this model may not be effective when too many zeros counts are encountered. More so, the threshold for modification when overdispersion was a serious cause of concern was not addressed. No contribution on the classes of overdispersion, whether mild, moderate, severe or not severe was addressed.

Shmuelli *et al* (2005) revived and explored Com-Poisson model. It is a two-parameter model, a convolution of three distributions: Poisson, Bernoulli, and Geometric distributions. The additional parameter is added to the distribution to model both under and overdispersion in count data. The application of the model and its usefulness to count data was discussed. Different three methods were used for parameter estimation simple weighted least squares, maximum likelihood method, and Bayesian approach method. The model was used to account for overdispersion which is usually associated with count data. Two examples of life studies were applied to validate the usefulness of the model. Both Poisson and Com-Poisson were used to model the count data. Poisson model was used to fit the retail of clothing; sample mean and sample variance were 3.56 and 11.31which indicates overdispersion. This model has not been fully explored and a comprehensive account of the statistical probabilities property remains a gap yet to be filled in the research. This model is limited in such that it is not viable to accommodate too frequent zero counts. Therefore, the literature on Com-Poisson is very scarce and the application is rarely come across in real life settings. Moreover, no point and threshold for classification of severity of overdispersion were discussed.

Famoye (2006) proposed Zero inflated Generalised Poisson (ZIPG) to model count data with excess zero. In the work zero-inflated Generalised Poisson regression model with application to domestic violence data, discovered that the conventional models could not model the count with a large number of zeros and Zero-inflated model proposed by Lambert (1992) was not adequate to model the data appropriately, hence for the need to propose ZIPG. The Maximum Likelihood method was used for parameter estimation. A score test is used to examine the adequacy of the model. The life study was modeled with ZIP, ZINB, and ZIPG but discovered that there was still high percentages proportion of zero with ZIP. Both ZINB and ZIPG were able to model the data sets appropriately but the iterative technique of ZINB did not converge which was also observed by Lambert (1992) when it was used to model the data sets.

Greene (2008) applied Negative Binomial Parameter (NBP) to model a large sample of German households study. The two variants of the Negative Binomial developed by Cameron & Trivedi (1986) posed problem. The problem of choosing the better model between the two models and carrying out appropriate statistical test and perform the parametric test for the non-nested form of the model between the models. Therefore, developed NBP an encompassing model that nested both Negative Binomial 1 (NB1) and Negative Binomial 11 (NB2), which enable the possibility of choosing the better and appropriate model between the two. But, the model fails to address the problem when overdispersion is of serious concern and the threshold for modification when severity is inherent in the count data.

Richards (2008) stated that in an Ecological study, overdispersion is often encountered which usually leads to the wrong conclusion. To overcome the problem; two approaches were examined. The first method employed compound distribution and the other method was the quasi approach which was to model the problem of overdispersion. Simulation studies were also used to demonstrate the two approaches. Both QAIC and AIC were used for model selection. In the interest for further studies simulation is suggested on choosing the better model.

Kimberly *et al* (2010) proposed a modified COM-Poisson regression model to address the problem of dispersion in count data. Poisson model was not adequate and efficient when used to model Airfreight breakage data. Both life and simulation studies were employed for the study. The simulation study was used to test for the accuracy estimation process. The model, its estimation and inference was discussed, and the relevance of the dispersion parameter was tested. Maximum likelihood method was used for parameter estimation. The model was compared with the other existing models for the count data, concluded that the new model is preferred to handle count data when there is a problem of overdispersion. One of the limitations of this model is that when encountered with excess zero the model may not be a good model for the data. An alternative model that can accommodate excess zero should be used.

Kimberly *et al* (2011) highlighted and stressed the point of difficulty often encountered in modeling count data with Poisson whenever the data are overdispersed. The limitation is usually experienced with the Poisson model as a result of equi-dispersion, that is, mean equal its variance which is very rare in a real life settings. COM-Poisson model is a flexible model with extra parameter capable to model count data whenever there is a problem of dispersion with the count data. Also, according to Shmueli (2005), Kimberly (2010), this model is a convolution of the three distributions which enable it to be more efficient to handle the problem of overdispersion. An extensive and comprehensive discussion of the different modified models and the application was later compared with existing models. The model was applied to life insurance. The limitation associated with this model is that on some occasions the predictive power may not portray the consistency property of the model and may not be adequate and appropriate to model when the scenario of excess zero arises in the study of count data.

Lee *et al* (2012) in the study" Analysis of overdispersed count data: Application to the Human papillomavirus infection in men study demonstrated that the Poisson model is not vibrant to model count data with overdispersion and excess zero count data. The Poisson, Negative binomial, Zero-inflated model, and Zero Negative Binomial model were used to model the data. In the study, discovered Poisson, Negative binomial and Zero-inflated Poisson models were not adequate to model the data but Zero-inflated Negative Binomial model was the only model adequate for the model which was equally observed by Famoye (2006). This study could not account for the problem of different classes of overdispersion

Alfonso *et al* (2013) provided another approach to handling overdispersion in count data when the Poisson model fails to account for overdispersion in count data. The standard error of the model should be considered to rectify the problem although the parameter estimates may be unbiased yet the parameter is under-estimated which can lead to a wrong and invalid conclusion of the model. The simulation method was used to study whether the increase of the sample sizes will lead to the reduction of the standard error or not. By increasing the sample sizes led to decrease of the estimate of standard error. Pearson's chi-square was used for testing of the dispersion parameter and non-parametric bootstrapped for simulation of the sample sizes but the method is limited in the sense that the simulation study could not give any account to the threshold of classification of overdispersion when it is severe.

Gsclobl (2013) presented Bayesian approach to count model estimation. Both Generalised and Negative Binomial regression to the model count data. Negative Binomial regression and the Generalized Poisson model could not account for the problem of excess zero. The data were modeled by Zero-inflated Generalised Poisson model to model the count data. Bayesian method was used for parameter estimation. There was no contribution on the count model when overdispersion is severe.

Ismail *et al* (2013) extended the work of Consul and Famoye (1992). Generalised Poisson 1(GP 1), Generalised Poisson 2(GP 2) and GP-P were developed, better models than Poisson and also GP-P, a model that nests parametrically both GP 1 and GP 2 regression model which makes it easier to carry out statistical tests and able to choose better model between the models. Oftentimes, in the application of Generalised Poisson, studies have shown that researchers did not know which model should fit the data set, when it comes to the use of Generalised Poisson. Exclusive work was carried out to demonstrate the appropriate model to fit the data sets. This study was limited in the sense that the threshold for the severity of overdispersion was not addressed.

Harrison (2014) signified that overdispersion may occur as a result of missing some important covariates, excess zero, and failure to account for such phenomenon can make parameter estimate biased which could lead to invalid or wrong conclusions. Observation level of random effects was considered to deal with the problem of overdispersion exhibits in count data. In the study, each of the observations coupled with its random effects was used to study the case of extra variation with Poisson. Studies have shown that observation level random effect has been a very good method in solving the problem of overdispersion but the method is very scarce, unpopular, and scanty research on it. Furthermore, the method is not capable to handle a problem of excess zero when it is used to model count with excess zeros.

Payne (2015) attempted to fill the gap by providing a comprehensive method and approach to solve the problem of overdispersion in count data. Different approaches were considered namely - the unadjusted Poisson regression, Deviance – scale adjusted Poisson regression, Pearson scale adjusted Poisson regression, Negative Binomial Regression, and two generalized linear mixed models (GLMM) were applied to model the count data. From the study, it was concluded that Negative binomial regression is preferred and appropriate to model after fitted first with Poisson model. In their views, different models that accommodate overdispersion should be considered for modeling the count data. The maximum likelihood method was used for the parameter estimation. Both simulation study and life study were considered.

Nwankwo and Godwin (2015) affirmed that the Negative binomial regression model addresses overdispersion in count data and most often Poisson regression provides a poor fit when used to model count data any time overdispersion is inherent in the data. In the study" Statistical model of Road traffic crashes data in Anambra State, Nigeria: A Poisson Regression Approach.   Three models were used to fit the data, namely, Poisson, Generalised Poisson, and Negative Binomial regression model. In their findings, they discovered that both the Poisson model and Generalised Poisson could not adequately handle the overdispersion presented in the data. Negative Binomial regression was able to model the data appropriately; AIC method was used for model selection and to show that the Negative Binomial regression is suitable for the Road crashes accident. However, this study did not reveal when the overdispersion is mild or severe and the point for modification was not provided.

Alaba *et al* (2017) used both Poisson and Negative Binomial to investigate the fertility pattern among women of childbearing age in Nigeria. Negative Binomial was used for modeling the data to account for overdispersion but the different classes of overdispersion were not discussed and there was no contribution on when overdispersion is of a serious cause of concern.

Endo (2020) studied the outbreak of Coronavirus 2019 (COVID-19) in the context of overdispersion both Poisson and Negative Binomial were used to model the data. The spread of the pandemic was considered and the data sets were overdispersed. Cases of both local and imported cases of infection of COVID-19 were collected from the World Health Organization (WHO). The samples used was made of quarantine and screened cases for infection of the virus. The count data was first fitted with the Poisson model to detect the presence of overdispersion before fitting with the alternative model which in this case is Negative Binomial. The Bayesian information method was used for the model selection. However, the level of overdispersion was not discussed whether the overdispersion was mild or severe also the threshold for classification was not mentioned.

Durmus and Guneri (2020) applied both the Poisson and GPR model to the number of strikes between 1984 and 2017 to account for the problem of overdispersion in case it occurs in count data. The problem with the common assumption of equality with Poisson regression usually cripples the Poisson model whenever overdispersion is inherent in the data. This problem of overdispersion may occur as a result of the inclusion of variables that are not associated with the study, and the dependence among the observation of the study. To remedy the problem, a distribution with an extra parameter that is suitable to capture the extra-variation and to rectify the malady would be used. Further, stated that GPR could be considered as an alternative model whenever this problem arises because this model has an extra parameter that can model the extra variation. Overdispersion was accounted for with GPR. The model adequacy was measured, opined that both AIC and BIC should be used for model selection when comparing model selection. Pearson statistics was used to detect the presence of the overdispersion, examined the ratio of the residual deviance to the degree of freedom of the Poisson model to indicate if the data is overdispersed or not. They discovered the data were overdispersed and eventually model fitted with GPR.

2.19 Literature Review on Threshold Methods.

Hansen (2000) developed a statistical point of view to threshold estimation. Despite the fact of the wide application of the study of threshold there was still limitation and underdeveloped use of the statistical application in estimation of the threshold. Least squares method was used in Time series analysis to propose an asymptotic method for constructing of intervals of the least threshold parameter of the interest. A recommendation was made on the use of simulation and bootstrap to determine threshold while considering the application of a threshold in study. This study also showed that a threshold is needed when there is a complexity in statistics to determine a criterion in a statistical study.

Hector and Gauthier (2003) presented and reviewed the extensive application of threshold in research studies. Several accounts and reviews of the application of threshold were discussed, relative to modeling, estimation, inference, and application to real life settings. The view which was in line with Hansen (2000) that estimation of threshold is still a gap yet to be filled in the study or research.

Theiwall *et al* (2005) considered three methods to propose a threshold for the impact of citation. The methods are arithmetic mean, geometric mean, and percentiles. The precision of these statistics was used to suggest which of the statistics can be used to propose a threshold for the study. The averaging method was recommended concerning the study of the impact of the citation. The method was also suggested and can be applied to other different fields of study. Poisson and Negative binomial models were also used to model the count data where Negative binomial was preferred to the Poisson model.

Mironov (2006) gave a comprehensive discussion on choosing threshold, explained how training set could be used to determine the threshold, described the method of rank statistic to determine threshold in the study of genomes in Bioinformatics. The rank statistic was employed to propose the threshold for the Site selection. In selecting the threshold, a value may be predefined to use as the yardstick to compare the threshold. Some random values were generated and their probabilities were calculated and this was compared to the probability of the value specified. Then both minimum and maximum values were determined.

Javidi and Mansoury (2017) applied Fuzzy set and Negative method to study gene selection of count data of Ant colony. Negative binomial is a flexible model that can model the count data when overdispersion is present in the data. A fuzzy set was also used for the classification of the gene into classes of which gene is redundant and which is not redundant. However, the classification of different classes of overdispersion when it is severe or not was not considered.

Bihn *et al* (2017) used Geometric mean as a tool to determine the threshold and criteria in studying the quality of Microbial water. Comprehensive outlines provided to determine the threshold for the quality of the water. In order to determine the threshold or criteria, a standard criterion or pre-defined values which will serve as the benchmark should exist and also a minimum and maximum threshold should be provided. In the study, twenty samples were used to carry out the test to propose the threshold for the quality of the water.

Marcos (2020) compared three common methods used in the estimation of threshold in ecotoxicology. The methods are maximum likelihood, Bayesian, and Piecewise regression. Root mean square error was used to estimate the accuracy of the method, and the ratio was used to show the relative improvement of the methodology and design. Among the three methods, Bayesian method was the preferred method for the study.

Yoonseok (2020) considered statistical inference to threshold regression estimation. Tests were carried out for the estimation of the regression parameter with restriction of the parameter of interest. Rank statistics were used to determine the threshold value of the selected observation. The simulation study was carried to establish the threshold values. The study was carried to study the threshold applied to the tipping point. In the study of the disposition of relocation of the white population whenever the threshold values of the minority threshold exceed the specified threshold. Through the study, it was revealed that most of the white re-located every ten decades when the threshold of the minority exceed the specified threshold.

Yuan *et al* (2015) on the account of using an averaging method to propose threshold, employed Averaging method and the Operative Characteristic curve (AUC) to the study of detection of breast cancer in medical research. These two methods were used to propose the threshold for detecting breast cancer and the risk factor for it. Comparing the two novel methods common in medical research discovered averaging method was more effective for proposing threshold in the medical research.

CHAPTER THREE

RESEARCH METHODOLOGY

3.1 Poisson Regression Model

Poisson model is a choice model for count data. Oftentimes, count data exhibit overdispersion whenever dealing with coun data. The model is modified when there may not be need for modification. This research studies the threshold for modification for some count data models which are Poisson, Negative Binomial, Com-Poisson, and Generalised Poisson model. Fuzzy set method and simulation study was used to classify the level of severity of overdispersion into four classes – not severe, moderately, severe, and very severe, and averaging method was used to determine the threshold. In this chapter, each of the methods will be discussed.

Considering the Poisson model,

$$P(Y = y) = \frac{e^{-\theta}\ \theta^{y}}{y!}, \qquad y = 0,1,2,3,\ldots \qquad \theta > 0 \tag{3.1}$$

$$Mean = E(X/Y) = \theta \tag{3.2}$$

$$Variance = Var(X/Y) = \theta \tag{3.3}$$

$$E(X/Y) = Var(X/Y) = \theta \tag{3.4}$$

The maximum likelihood method is used for the parameter estimation, usually, the log likelihood function is obtained by taking the log-likelihood function thus

$$\ln f(y/\theta) = y\ln(\theta) - \theta - \ln(y!) \tag{3.5}$$

Such that

$$\ln(\theta) = \sum_{i=0}^{n} x_i\beta = \quad \text{with } x_0 = 1 \tag{3.6}$$

$$Mean = E(X) = \mu \tag{3.7}$$

$$Variance = E(X) = \mu \tag{3.8}$$

Given

$$\lambda = |v - m| = 0 \tag{3.9}$$

Such that when

$$\lambda = 0 \quad \text{there is no dispersion} \tag{3.10}$$

$$\lambda \prec 0 \quad \text{there is underdispersion} \tag{3.11}$$

$$\lambda \succ 0 \quad \text{there is overdispersion} \tag{3.12}$$

where *m* is the mean and *v* is the variance, λ is the degree of severity of overdispersion.

3.2 Fuzzy Set Approach

A fuzzy set is used whenever classification lacks boundaries and criteria; unlike a classical set that element can either belong or not belong to a set. In a fuzzy set, an element can either belong or partially or not belong to a set with the use of membership function which shows the degree of belonging of the element to the set.

Consider *S*, the fuzzy subset of *D*, which is characterised by the membership function of *d* in the interval [0 , 1] $\mu_s(D): D \rightarrow [0.1]$ where $\mu_s(D)$ is the degree to which *D* belongs to S.

In general,

$$S = \{D, \mu_S(D)\} \tag{3.13}$$

3.3 Fuzzy-C-Partition

Fuzzy-c- Partition

The fuzzy-c-Partition method is used when dealing with heterogeneous members that need to be classified into homogenous groups or classes where there will be a homogenous member within the classes. In the research, this method is used.

Consider universal set of heterogeneous members called *D* such that

$$D = \{d_1, d_2, ...\} \tag{3.14}$$

Using Fuzzy –c- partition for the classification

$$D = \{(d_{ji}, .. d_{jn} ... d_{jn}, ... d_{jn})\} \quad \text{, where j} = 1,2,...\text{c, i} = 1,2,...\text{n} \tag{3.15}$$

That is *c*- classes with their respective elements. Each of the classes has a membership function that shows the degree of belonging to the subset given as

$$\mu_s(D) = \begin{cases} 1 & if\ D \text{ is totally in c} \\ 0 < \mu_s(D) < 1 & if\ D \text{ is partially in c} \\ 0 & if\ D \text{ is not in c} \end{cases} \tag{3.16}$$

On Overdispersion

Now, consider overdispersion, using Fuzzy c- partition for different overdispersion levels which are classified as not severe, moderately severe , severe, and very severe, that is universal set of D into different classes overdispersion.

Given universal set *D* such as

$$D = \{d_1, d_2, ... d_n\} \quad \text{where} \quad 0 \le d \le 100 \tag{3.17}$$

Here, *D* is the universal set of overdispersion percentages and *d* is the element in each class of the set. Therefore, *D* was calculated as

$$D = \left(\frac{(v-m)}{m} \right) \% \tag{3.18}$$

categorised as

$$D = \{ d \le p, \mathrm{p} < \mathrm{d} \le \mathrm{r}, \mathrm{r} < \mathrm{d} \le \mathrm{w}, \ \mathrm{d} \ge \mathrm{w} \} \tag{3.19}$$

where *p, r, w* are the predefined thresholds for the study given as *p=10%, r=40%,* and *w=70%.* A membership function is a basic tool for a fuzzy set. It shows the degree of closeness of belongingness to a particular subset of the universal set. Therefore, we constructed a membership function that suits our study which is given below;

$$\mu_\lambda(D) = \begin{cases} 1 & if\ D \text{ is totally in } \lambda \\ 0 < \mu_\lambda(D) < 1 & if\ D \text{ is partially in } \lambda \\ 0 & if\ D \text{ is not in } \lambda \end{cases} \tag{3.20}$$

where λ is the set of severity that is, not severe, moderately severe, severe and very severe.

Considered Bezdek (1981) and Yang (1993) for the classification of overdispersion into four distinct classes with their respective membership function.

$$\mu_\lambda(D) = \begin{cases} 0 & if\ d \le p & \text{(not severe)} \\ \frac{d-p}{r-p} & if\ p < d \le r & \text{(moderately severe)} \\ \frac{d-r}{w-r} & if\ r < d < w & \text{(severe)} \\ 1 & if\ d \ge w & \text{(very severe)} \end{cases} \tag{3.21}$$

3.4 Fuzzy Class Poisson Regression

Given a mixture of distribution,

$$g(y/\lambda) = \sum_{d}^{D} \phi\, g_d(y/\lambda_d) \qquad \sum \phi = 1 \tag{3.22}$$

Yang (1993) considered a Fuzzy Class Maximum Likelihood (FCML) such that if θ is replaced by fuzzy class variable λ_d and $g_d(y/\lambda_d)$ by in place of f(y/θ) then, fuzzy class regression function is

$$G(\lambda, d, \beta) = \sum_{i}^{c} \sum_{1}^{p} \lambda_d \ln f(y_i / \theta) + w \sum_{i}^{c} \sum_{i}^{p} \lambda_d \ln d_i \qquad (3.23)$$

where

$$\ln \lambda_d = \beta_{0d} + \sum_{i=1}^{p} x_{i1} \beta_{1d} \qquad (3.24)$$

The derivaties of $G_{r,y}(\lambda,d,\beta)$ with respect to β

$$\frac{\partial}{\partial \beta_{1d}} G_{r,v}(\lambda, d, \beta) = \sum_{i=1}^{n} \lambda_d \frac{\partial \ln g_d(y/\lambda_d)}{\partial \lambda_d} \frac{\partial \lambda_{\mathrm{d}}}{\partial \beta_{1\mathrm{d}}} \qquad (3.25)$$

where

$$\frac{d \ln g_d(y/\lambda_d)}{\partial \lambda_d} = \left(\frac{y_i}{\lambda_d} - 1 \right) \qquad (3.26)$$

$$\frac{\partial \lambda_d}{\partial \beta_{1d}} = x_{il} \lambda_d \qquad (3.27)$$

Thus, we have

$$\frac{\partial}{\partial \beta_{id}} G_{r,y}(\lambda, d, \beta) = \sum_{i=1}^{n} \lambda_d^r (y_i - \lambda_d) x_{il} \qquad (3.28)$$

Estimating the parameters by maximizing the function

$$G_{r,y}(\lambda,d,\beta) \qquad (3.29)$$

with restrictions

$$\sum_{d=1}^{D} \lambda_d = 1 \qquad (3.30)$$

Consider,

$G_{r,y}(\lambda, d, \beta)$ (3.31)

Taking the derivatives of (3.31) with respect to D and λ_d respectively we have

$$\hat{D} = \frac{n\sum_{i=1}\lambda_d^r}{\sum_{i=1}^{n}\sum_{s=1}^{D}\lambda_{sd}^r} \tag{3.32}$$

$$\hat{\lambda}_d = \left(\sum_{q=1}^{D} \frac{\ln g_d(y/\lambda_d) + w\ln \hat{D})^{\frac{1}{\hat{m}-1}}}{\ln g_d(y/(y/\lambda_{i,q})\lambda_d) + w\ln \hat{D})^{\frac{1}{\hat{m}-1}}} \right) \tag{3.33}$$

i=1,2,…,n d=1,2,…,D

3.5 Negative Binomial Distribution

The probability mass function of a Negative Binomial Distribution is

$$\Pr(Y = k) = \binom{k + r - 1}{r - 1}(1 - p)^k p^r \tag{3.34}$$

where r is the number of successes, k is the number of failures and p is the probability of success

$$E(Y) = \frac{pr}{1 - p} \tag{3.35}$$

$$Var(Y) = \frac{pr}{(1 - p)^2} \tag{3.36}$$

3.6 Fuzzy Class Negative Binomial Regression

Also, for Negative Binomial replace θ by fuzzy class variable λ_d and $g_d(y/\lambda_d)$ in place of $h_y(y_i, \theta, \alpha)$ such that the fuzzy class regression function is

$$H_y(\lambda, d, \beta) = \sum_i^c \sum_1^p \lambda_d \ln h_y(y_i / \alpha) + w \sum_i^c \sum_i^p \lambda_d \ln d_i \tag{3.37}$$

where

$$h_y(y, \lambda_d \beta) = \frac{\Gamma(y+\alpha)}{\Gamma\alpha \text{ y!}} \frac{\lambda^y \alpha}{(\lambda_d + \alpha)^{y+\alpha}} \tag{3.38}$$

The mean is

$$E(y) = \lambda_d \tag{3.39}$$

And

$$Var(y) = \lambda_d (1 + \lambda_d) / \alpha)$$

$$Var(y) = \lambda_d ((1 + \lambda_d) / \phi) \tag{3.40}$$

The variance is quadratic in mean. The Negative binomial distribution can also be modeled by the dispersion parameter (θ) such as

$$\alpha = 1/\theta \;\; \theta = 1/\phi \tag{3.41}$$

When θ greater than one, when $\alpha \to 0$, the Negative Binomial distribution is reduced to standard Poisson distribution with parameter λ. Parameter estimation of β and λ is done by maximizing the log likelihood function

$$l(\beta, \lambda_d, y) = \sum_d^D \log\left(\frac{\Gamma(y+\theta^{-1})}{\Gamma(1-(y+\theta^{-1}))} \right) - \log(y!) - (y+\theta^{-1}) \log(1+\theta\lambda_d) + \log(y\lambda) + yx^T \beta \tag{3.42}$$

3.7 Com-Poisson Regression Model

The Com-Poisson is defined to be distributed with Probability Mass Function (PMF) given by

$$P(Y = y) = f(y, \theta, v) = \frac{\theta^y}{(y!) Z(\theta, v)} \tag{3.43}$$

where the normalizing constant Z,

$$Z = (\theta, v) = \sum_{j=0=}^{\infty} \frac{\theta^{j}}{(j)!^{y}} \tag{3.44}$$

where v is the extra parameter, as v=0, Com-Poisson becomes geometric distribution when v =1, Com-Poisson becomes Poisson distribution and when v= ∞ it becomes Bernoulli trial.

$$Mean = \sum_{J=0}^{\infty} \frac{j\theta^{j}}{(j!)^{v} Z(\theta, v)} \tag{3.45}$$

$$Variance = \sum_{j=0}^{\infty} \frac{j^{2}\theta^{j}}{(j!)^{v} Z(\theta, v)} - \left(\sum_{j=0}^{\infty} \frac{j^{2}\theta^{j}}{(j!)^{v} Z(\theta, v)} \right)^{2} \tag{3.46}$$

3.8 Fuzzy Class Com-Poisson Regression Model

Also, Com-Poisson replace θ by fuzzy class variable λ_d and $g_d(y/\lambda_d)$ in place of $f_y(y_i, \theta, v)$ such that the fuzzy class regression function is

$$G(\lambda, d, \beta) = \sum_{i}^{c} \sum_{1}^{p} \lambda_{d} \ln f(y_{i} / v) + w \sum_{i}^{c} \sum_{i}^{p} \lambda_{d} \ln d_{i} \tag{3.47}$$

where

$$P(Y = y) = f(y, \theta_{d}, v) = \frac{\theta_{d}^{y}}{(y!) Z(\theta_{d}, v)} \tag{3.48}$$

Mean and Variance are given as

$$Mean = \sum_{J=0}^{\infty} \frac{j\theta_{d}^{j}}{(j!)^{v} Z(\theta_{d}^{j}, v)} \tag{3.49}$$

$$Variance = \sum_{j=0}^{\infty} \frac{j^{2}\theta_{d}^{j}}{(j!)^{v} Z(\theta_{d}^{j}, v)} - \left(\sum_{j=0}^{\infty} \frac{j\theta_{d}^{j}}{(j!)^{v} Z(\theta_{d}^{j}, v)} \right)^{2} \tag{3.50}$$

$$\log L_{i}(\lambda_{d}, v / y_{i}) = y_{i} \log \lambda_{d} - v \log y! - \log z(\lambda_{d}, v) \tag{3.51}$$

Summing over n, the log likelihood is given by

$$\log = \sum_{i=1}^{n} y_i \log \lambda_d - v_i \sum y_i \log y_i ! - \sum_{i=1}^{n} \log z(\lambda_d, v) \quad (3.52)$$

3.9 Generalised Poisson Model

A random variable Y is said to have Generalised Poisson distribution with parameters α >0, θ>0 denoted by G(α, θ) if the probability function is given as

$$P(Y = y) = \begin{cases} \alpha[(\alpha + y\theta) + \theta y)y - 1]\frac{1}{y!}e^{(-\alpha - y\theta)} \\ 0 \end{cases} \quad , \quad y = 0,1,2,\ldots \quad (3.53)$$

Mean and Variance of the distribution given as

$$E(y/\alpha,\theta) = \frac{\alpha}{1-\theta} \quad \text{and} \quad Var(y/\alpha,\theta) = \frac{\alpha}{(1-\theta)^3} \quad (3.54)$$

3.10 Fuzzy Class Generalised Poisson Regression Model

Likewise, replace α by fuzzy class variable λ_d and $g_d(y/\lambda_d)$ in place $P(Y = y_i / \alpha, \theta)$

such that the fuzzy class regression function is

$$G(\lambda, d, \beta) = \sum_{i}^{c} \sum_{1}^{p} \lambda_d \ln g_y (y_i / \lambda_d, \theta) + w \sum_{i}^{c} \sum_{i}^{p} \lambda_d \ln y_{i_i} \quad (3.55)$$

,

where

$$g_d(y/\lambda_d,\theta) = \begin{cases} \lambda_d[(\lambda_d + y\theta) + \theta y)y - 1]\frac{1}{y!}e^{(-\lambda_d - y\theta)} \\ 0 \end{cases} \quad y = 0,1,2\ldots\ldots \quad (3.56)$$

With

$$E(y/\lambda_d,\theta)=\frac{\lambda_d}{1-\theta} \quad \text{and} \quad Var(y/\lambda_d,\theta)=\frac{\lambda_d}{(1-\theta)^3} \tag{3.57}$$

3.11 Monte Carlo Design of Experiment of the Study

Data Generation Process

Four different models namely Poisson, Negative Binomial, Com-Poisson and Generalised Poisson models were used for the simulation study. Different data sets were generated for overdispersion to study the case of severity of overdispersion in count data. Membership functions were also constructed for classifying the dispersion on the basis of severity, where average threshold is

$$Average=\frac{\sum_{i=1}^{n} D_i}{n} \tag{3.58}$$

where i=1,2,3,...n

D_i is the different dispersion percentages and n is the total number of observations. Data were simulated using Monte Carlo design of experiment to establish the presence of overdispersion in the count data. For each of model; varying means and variances were generated and dispersion percentages were computed from 3.61. Some predefined thresholds were selected such that

(p=10%, r= 40%, w =70%) (3.59)

D = (not severe, moderately severe, severe, very severe) (3.60)

Element of the Universal set D is computed by

$$D=\left(\frac{v-m}{m}\right)\% \tag{3.61}$$

Define the membership function as

$$\mu_{\lambda}(D) = \begin{cases} 0 & if\ d \le 10 & \text{(not severe)} \\ \dfrac{4d-40}{210} & if\ 10 < d \le 40 & \text{(moderately severe)} \\ \dfrac{d}{70} & if\ 40 < d \le 70 & \text{(severe)} \\ 1 & if\ d \ge 70 & \text{(very severe)} \end{cases} \tag{3.62}$$

Simulated for n=20, 30, 50, 100, 200, 300, 500, 1000, 5000

CHAPTER FOUR

RESULTS AND DISCUSSION

4.0 Chapter Overview

This chapter presents the result of the simulation study and the application to life study. The simulation study will first be discussed before the life study. A simulation study was carried out to determine the threshold for the point of modification for each of the models namely: Poisson (PO), Negative Binomial (NB), Com-Poisson (CP), and Generalised Poisson model (GP). The averaging method was used to determine the threshold. The results are presented in table 4.1-4.24, where the membership function is in parentheses.

4.1 Fuzzy Set Classification of Different Levels of Overdispersion.

Table 4.1-4.5 show the results for the Fuzzy c- partition simulation for Poisson model; Table 4.1 shows when n=20, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01 \, and \, \sigma^2 = 0.05,$,D={1.48(0), 8.90(0), 12.95(0.056), 39.02(0.553), 98.84(1)}. For $\mu = 0.05 \, and \, \sigma^2 = 0.1,$ D={0.31(0), 9.52(0), 10.27(0.005), 39.92(0.976), 40.36(0.577), 65.65(0.937)}.For $\mu = 0.5 \, and \, \sigma^2 = 0.55$ D={0,(9.52(0), 0.11(0.002),34.54(0.467), 53.86(0.769)}, For $\mu = 1 \, \sigma^2 = 1.5,$ D={(2.23(0), 9.31(0), 11.92(0.037), 33.94(0.457),42.42(0.606),68.35(0.976)}For $\mu = 2 \, and \, \sigma^2 = \{0.22(0), 9.30(0), 2.5,$ D =13.52(0.067),9.92(0.976),42.90(0.613),56.11(0.802), 117.11(1)}. For $\mu = 10 \, \sigma^2 = 10.5,$ D={5.56(0), 10.16(0.003), 39.63(0.564), 40.41(0.577), 53.68(0.767), 83.03(1), 116(1).} For $\mu = 50 \, \sigma^2 = 55,$ D={1.72(0), 9.52(0), 10.14(0.003), 33.97(0.457), 44.46(0.664), 63.79(0.911)}, For $\mu = 100 \, and \, \sigma^2 = 105,$ D={1.13(0), 9.81(0), 10.77(0.015), 36.10(0.497), 69.72(0.996), 96.99(1)} respectively.

Table 4.1. Fuzzy set classification of different levels of overdispersion of (PO) n=20,30.

n=20

		Not Severe		Moderate-ly severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	1.48 (0)	8.90 (0)	12.95 (0.056)	39.02 (0.553)	-	-	98.84 (1)	-
0.05	0.1	0.31 (0)	9.52 (0)	10.27 (0.005)	39.92 (0.570)	40.36 (0.577)	65.65 (0.937)	-	-
0.5	0.55	0	9.52 (0)	10.11 (0.002)	34.54 (0.467)	53.86 (0.769)	-	-	-
1	1.5	2.23 (0)	9.31 (0)	11.92 (0.037)	33.94 (0.457)	42.42 (0.606)	68.35 (0.976)	-	-
2	2.5	0.22 (0)	9.30 (0)	13.52 (0.067)	39.92 (0.570)	42.90 (0.613)	56.11 (0.802)	117.11 (1)	-
10	10.5	5.56 (0)	-	10.16 (0.003)	39.63 (0.564)	40.41 (0.577)	53.68 (0.767)	83.03 (1)	116.34 (1)
50	55	1.72 (0)	9.52 (0)	10.14 (0.003)	33.97 (0.457)	44.46 (0.664)	63.79 (0.911)	-	-
100	105	1.13 (0)	9.81 (0)	10.77 (0.015)	36.10 (0.497)	69.72 (0.996)	-	96.99 (1)	-
				n=30					
		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	1.48 (0)	7.52 (0)	14.40 (0.084)	39.02 (0.553)	61.70 (0.881)	65.87 (0.941)	73.75 (1)	-
0.05	0.1	1.72 (0)	14.40 (0)	11.45 (0.028)	39.02 (0.553)	46.15 (0.659)	-	71.05 (1)	72.73 (1)
0.5	0.55	0	5.56 (0)	10.11 (0.002)	35.71 (0.490)	47.71 (0.862)	65.60 (0.937)	-	-
1	1.5	0	8.90 (0)	13.32 (0.063)	35.04 (0.477)	43.69 (0.624)	54.65 (0.781)	79.78 (1)	225.50 (1)
2	2.5	0	9.52 (0)	11.44 (0.028)	37.89 (0.531)	43.69 (0.624)	58.33 (0.833)	96.55 (1)	147.83 (1)
10	10.5	0.84 (0)	8.06 (0)	11.76 (0.034)	39.92 (0.976)	40.03 (0.004)	61.70 (0.881)	-	-
50	55	0	8.50 (0)	10.77 (0.015)	33.11 (0.440)	46.27 (0.661)	51.68 (0.738)	71.05 (1)	124.16 (1)
100	105	2.84 (0)	8.03 (0)	11.45 (0.028)	32.65 (0.431)	45.85 (0.655)	68.35 (0.976)	83.32 (1)	111.11 (1)

NOTE: The value in the bracket is the membership value.

Table 4.1 when n=30 shows the results for the Fuzzy c- partition simulation when n=30, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01, \sigma^2 = 0.05,$,D={1.48(0), 7.52 (0), 14.40(0.084), 39.02(0.553), 61.70(0.881), 65.87(0.941), 73.75(1) }. For $\mu = 0.05, \sigma^2 = 0.1$, D={1.72(0), 14.40 (0), 11.45 (0.028), 39.02(0.553), 46.15 (0.659).71.05(1), 72.73(1)}. For $\mu = 0.5, \sigma^2 = 0.55$ D= {0, 5.56(0), 10.11(0.002), 35.71(0.490), 47.71(0.862), 65.60(0.937)}. For $\mu = 1$ $\sigma^2 = 1.5$, D={(0, 8.90(0), 13.32(0.0063), 35.04(0.477), 43.69(0.624), 54.65(0.781), 79.78(1), 225.50(1)}. For $\mu = 2$ $\sigma^2 = 2.5$, D={0, 9.52(0), 11.44 (0.028), 37.89(0.531), 43.69(0.624) 58.33(0.833), 99.55(1), 147.83(1)}. For $\mu = 10$ $\sigma^2 = 10.5$, D={0.84(0), 0.86(0), 11.76(0.0.034), 39.92(0.976), 40.03(0.004), 61.70(0.881)} For $\mu = 50$ $\sigma^2 = 55$, D={ 0, 8.50(0), 10.77(0.015), 33.11 (0.440), 46.27(0.661), 51.68(0.738), 71.05(1), 124.16(1), For $\mu = 100$ $\sigma^2 = 105$, D={2.84 (0), 9.81(0), 8.33 (0), 11.45(0.028), 32.65(0.431), 45.45(0.655), 68.35(0.976), 83.21(1), 111.11(1) } respectively.

Table 4.2 shows the results for the Fuzzy c- partition simulation when n=50, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01, \sigma^2 = 0.05,$,D={0, 8.50 (0), 10..47(0.009), 35.43 (0.484), 47.41 (0.667), 65.81(0.940), 78.13 (1), 189.13(1) }. For $\mu = 0.05, \sigma^2 = 0.1$, D={1.22(0), 8.50 (0), 11.45 (0.028), 37.84 (0.530), 42.65 (0.609). 61.44(0.878), 87.32(1), 97.17(1)}. For $\mu = 0.5, \sigma^2 = 0.55$ D= {0.25(0), 9.81 (0), 11.45 (0.028), 39.51(0.562), 40.57 (0.580), 67.65 (0.966, 82.3291), 111.11(1))}. For $\mu = 1$ $\sigma^2 = 1.5$, D={(0, 8.90(0), 10.47 (0.009), 39.02 (0.533), 61.70 (0.941), 65.876(0.941), 73.75(1), 189.13(1)}. For $\mu = 2$ $\sigma^2 = 2.5$, D={1.46(0), 9.52(0), 10.38 (0.007), 31.50(0.410), 40.57(0.580), 68.35(0.976), 71.05(1), 142.55(1)}. For $\mu = 10$ $\sigma^2 = 10.5$, D={0, 9.62(0), 10.47(0.009), 39.36(0.559), 46.15(0.659), 67.65(0.966), 126.19(1)} For $\mu = 50$ $\sigma^2 = 55$, D={ 0.27(0), 9.52(0), 10.42(0.008), 35.71(0.490), 43.08(0.615), 69.72(0.996), 71.05(1), 185.84(1)}.For $\mu = 100$ $\sigma^2 = 105$, D={0, 9.52(0), 10.11(0.002), 40.00(0.571), 43.69(0.624), 61.70(0.881), 87.32(1)} respectively

Table 4.2. Fuzzy set classification of different levels of overdispersion of (PO) n=50, 100.

n=50

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	0	9.81 (0)	10.47 (0.009)	35.43 (0.484)	47.41 (0.667)	65.81 (0.940)	78.13 (1)	189.13 (1)
0.05	0.1	0.22 (0)	8.50 (0)	11.45 (0.028)	37.84 (0.530)	42.65 (0.609)	61.44 (0.878)	87.32 (1)	97.17 (1)
0.5	0.55	0.25 (0)	9.81 (0)	11.45 (0.028)	39.51 (0.562)	40.57 (0.580)	67.65 (0.966)	82.32 (1)	111.11 (1)
1	1.5	0	8.90 (0)	10.47 (0.009)	39.02 (0.553)	61.70 (0.881)	65.87 (0.941)	73.75 (1)	189.13 (1)
2	2.5	1.46 (0)	9.52 (0)	10.38 (0.007)	31.50 (0.410)	40.57 (0.580)	68.35 (0.976)	71.05 (1)	142.55 (1)
10	10.5	0	9.62 (0)	10.47 (0.009)	39.36 (0.559)	46.15 (0.659)	67.65 (0.966)	126.19 (1)	-
50	55	0.27 (0)	9.52 (0)	10.42 (0.008)	35.71 (0.490)	43.08 (0.615)	69.72 (0.996)	71.05 (1)	185.84 (1)
100	105	0	9.52 (0)	10.11 (0.002)	40.00 (0.571)	43.69 (0.624)	61.70 (0.881)	84.33 (1)	87.32 (1)
		n=100							
		Not Severe		Moderate-ly		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	0	9.81 (0)	10.38 (0.007)	39.92 (0.976)	40.16 (0.574)	68.35 (0.976)	82.32 (1)	116.34 (1)
0.05	0.1	0.27(0)	9.52 (0)	10.47 (0.009)	39.02 (0.553)	41.04 (0.586)	68.35 (0.976)	71.05 (1)	119.23 (1)
0.5	0.55	0.27(0)	9.52 (0)	11.44 (0.028)	40.00 (0.571)	41.95 (0.599)	62.54 (0.893)	76.85 (1)	246.05 (1)
1	1.5	0	9.52 (0)	10.47 (0.009)	37.17 (0.518)	41.49 (0.593)	69.72 (0.996)	82.32 (1)	112.39 (1)
2	2.5	0	9.62 (0)	10.24(0.004)	39.87 (0.569)	42.20 (0.603)	69.00 (0.986)	72.73 (1)	82.32 (1)
10	10.5	0.22(0)	9.81 (0)	10.38 (0.007)	39.92 (0.570)	41.59 (0.594)	60.39 (0.863)	97.17 (1)	149.80 (1)
50	55	0	9.62 (0)	10.77 (0.015)	40.00 (0.571)	40.03 (0.572)	69.72 (0.996)	71.05 (1)	171.43 (1)
100	105	0.25(0)	9.52 (0)	10.24 (0.004)	39.02 (0.553)	43.00 (0.614)	66.79 (0.954)	70.15 (1)	89.49 (1)

Table 4.2 shows the results for the Fuzzy c- partition simulation when n=100, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01, \sigma^2 = 0.05,$,D={0, 9.81(0), 10.38(0.007), 39.92(0.976), 40.16(0.574), 68.35(0.976), 82.32(1), 116.34(1) }. For $\mu = 0.05, \sigma^2 = 0.1$, D={0.27(0), 9.52(0), 10.47(0.009), 39.02(0.553), 41.04(0.586), 68.35(0.976), 71.05(1). 119.23(1)}.. For $\mu = 0.5, \sigma^2 = 0.55$ D= {0.27(0), 9.52(0), 11.44(0.028), 40.00(0.571), 41.95(0.599), 62.54(0.893), 76.85(1), 246.05(1)}. For $\mu = 1$ $\sigma^2 = 1.5$, D={(0, 9.52(0), 10.47 (0.009), 37.17(0.518), 41.49(0.593), 69.72(0.996), 82.32(1), 112.39(1)}. For $\mu = 2$ $\sigma^2 = 2.5$, D={0, 9.62(0), 10.24(0.004), 39.87(0.569), 42.20(0.603), 69.00(0.986), 72.73(1), 82.32(1)}. For $\mu = 10$ $\sigma^2 = 10.5$, D={0.22(0), 9.81(0), 10.38(0.007), 39.92(0.570), 41.59(0.594), 60.39(0.863), 97.17(1), 149.80(1)}. For $\mu = 50$ $\sigma^2 = 55$, D={ 0, 9.62(0), 10.77(0.015), 40.00(0.571), 40.03(0.572), 69.72(0.996). 71.05(1), 171.43(1)}. For $\mu = 100$ $\sigma^2 = 105$, D={0, 9.52(0), 10.24(0.004), 39.02(0.553), 43.00(0.614), 66.79(0.954), 70.15(1), 89.49(1)} respectively.

Table 4.3 shows the results for the Fuzzy c- partition simulation when n=200, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01, \sigma^2 = 0.05,$,D={0, 9.62(0), 10.11(0.002), 40.00 (0.571), 40.57(0.580), 69.21(0.996), 71.05(1), 185.84(1)}. For $\mu = 0.05, \sigma^2 = 0.1$, D={0, 9.52(0), 10.24(0.004), 40.00 (0.571), 40.34(0.576), 69.00(0.986), 72.73(1). 116.34(1)}. For $\mu = 0.5, \sigma^2 = 0.55$ D= {0, 9.52(0), 10.77(0.015), 39.36(0.599), 40.41(0.577), 69.72(0.996), 71.05(1),141.82(1)}. For $\mu = 1, \sigma^2 = 1.5$ D= {0, 9.52(0), 10.77(0.015), 39.36(0.599), 40.41(0.577), 69.72(0.996), 71.05(1),141.82(1)}. For $\mu = 2$ $\sigma^2 = 2.5$, D={0.22(0), 9.81(0), 10.24(0.004), 39.922(0.976), 41.14 (0.588), 67.65(0.966), 70.15(1), 189.13(1)}. For $\mu = 10$ $\sigma^2 = 10.5$, D={0.25(0), 9.85(0), 10.11(0.002), 40.00 (0.571), 41.14(0.588), 66.79(0.954), 70.15(1), 180.08(1)}. For $\mu = 50$ $\sigma^2 = 55$, D={ 0, 9.81(0), 10.11(0.002), 37.48 (0.523), 40.16(0.574), 68.35(0.976). 77.47(1), 159.09(1)}. For $\mu = 100$ $\sigma^2 = 105$, D={0, 9.81(0), 10.11(0.002), 40.00 (0.571), 41.27(0.590), 69.00(0.986), 76.85(1), 210.91(1)} respectively.

Table 4.3 Fuzzy set classification of different levels of overdispersion of (PO) n=200,300.

n=200

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	0	9.62 (0)	10.11 (0.002)	40.00 (0.571)	40.57 (0.580)	69.72 (0.996)	71.05 (1)	185.84 (1)
0.05	0.1	0	9.52 (0)	10.24 (0.004)	40.00 (0.571)	40.34 (0.576)	69.00 (0.986)	72.73 (1)	116.34 (1)
0.5	0.55	0	9.52 (0)	10.77 (0.015)	39.36 (0.559)	40.41 (0.577)	69.72 (0.996)	71.05 (1)	141.82 (1)
1	1.5	0	9.52 (0)	10.77 (0.015)	39.36 (0.559)	40.41 (0.577)	69.72 (0.996)	71.05 (1)	141.82 (1)
2	2.5	0.22 (0)	9.81 (0)	10.24 (0.004)	39.92 (0.976)	41.14 (0.588)	67.65 (0.966)	70.15 (1)	189.13 (1)
10	10.5	0.25 (0)	9.85 (0)	10.11 (0.002)	40.00 (0.571)	41.14 (0.588)	66.79 (0.954)	70.15 (1)	180.08 (1)
50	55	0	9.81 (0)	10.11 (0.002)	37.48 (0.523)	40.16 (0.574)	68.35 (0.976)	77.47 (1)	159.09 (1)
100	105	0.25 (0)	9.81 (0)	10.11 (0.002)	40.00 (0.571)	41.27 (0.590)	69.00 (0.986)	76.85 (1)	210.91 (1)
		n=300							
		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	0	9.81 (0)	10.24 (0.004)	40.00 (0.571)	40.03 (0.572)	69.72 (0.996)	71.05 (1)	171.43 (1)
0.05	0.1	0	9.81 (0)	10.38 (0.007)	39.92 (0.976)	40.16 (0.574)	69.72 (0.996)	71.05 (1)	189.13 (1)
0.5	0.55	0	9.62 (0)	10.11 (0.002)	39.92 (0.976)	40.21 (0.574)	69.72 (0.996)	75.93 (1)	233.18 (1)
1	1.5	0	9.62 (0)	10.24 (0.004)	40.00 (0.571)	40.34 (0.576)	68.35 (0.976)	71.05 (1)	259.46 (1)
2	2.5	0	9.81 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.004)	69.00 (0.986)	70.15 (1)	171.43 (1)
10	10.5	0	9.62 (0)	10.11 (0.002)	39.92 (0.976)	40.57 (0.580)	67.36 (0.962)	71.05 (1)	216.67 (1)
50	55	0	9.52 (0)	10.11 (0.002)	40.00 (0.571)	40.42 (0.577)	69.72 (0.996)	72.73 (1)	169.75 (1)
100	105	0	9.81 (0)	10.27 (0.005)	39.92 (0.976)	40.11 (0.573)	67.65 (0.966)	71.05 (1)	216.67 (1)

Table 4.3 shows the results for the Fuzzy c- partition simulation for when n=300, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01, \sigma^2 = 0.05,$,D={0, 9.81(0), 10.24(0.004), 40.00 (0.571), 40.03(0.572), 69.72(0.996), 71.05(1), 171.43(1)}. For $\mu = 0.05, \sigma^2 = 0.1,$ D= {0, 9.81(0), 10.38(0.007), 40.16 (0.574), 69.72 (0.996), 71.05(1). 189.13(1)}. For $\mu = 0.5, \sigma^2 = 0.55$ D= {0, 9.62(0), 10.11(0.002), 39.92(0.976), 40.21(0.574), 69.72(0.996), 75.93(1), 233.18(1)}. For $\mu = 1, \sigma^2 = 1.5$ D= {0, 9.62(0), 10.24(0.004), 40.00 (0.571), 40.34(0.576), 68.35(0.976), 71.05(1), 259.46(1)}. For $\mu = 2$ $\sigma^2 = 2.5,$ D= {0, 9.81(0), 10.11(0.002), 40.00 (0.571), 40.03 (0.004), 69.00(0.986), 70.15(1), 171.43(1)}. For $\mu = 10$ $\sigma^2 = 10.5,$ D= {0, 9.62(0), 10.11(0.002), 39.92 (0.976), 40.57(0.580), 67.36(0.962), 71.05(1), 216.67(1)}. For $\mu = 50$ $\sigma^2 = 55,$ D= {0, 9.52(0), 10.11(0.002), 40.00 (0.571), 40.42(0.577), 69.72(0.996). 72.73(1), 169.75(1)}. For $\mu = 100$ $\sigma^2 = 105,$ D= {0, 9.81(0), 10.27(0.005), 39.92 (0.976), 41.11 (0.573), 67.65(0.966), 71.05(1), 261.67(1)} respectively.

Table 4.4 shows the results for the Fuzzy c- partition simulation for when n=500, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01, \sigma^2 = 0.05,$,D={0, 9.81(0), 10.24(0.004), 40.00 (0.571), 40.03(0.572), 69.72(0.996), 71.05(1), 246.05(1)}. For $\mu = 0.05, \sigma^2 = 0.1,$ D= {0, 9.81(0), 10.24(0.004), 39.92(0.976) 40.41 (0.577), 69.72 (0.996), 70.15(1). 151.60(1)}. For $\mu = 0.5, \sigma^2 = 0.55$ D= {0, 9.85(0), 10.16(0.003), 40.00 (0.571), 40.16(0.574), 69.72(0.996), 70.51(1), 245.45(1)}. For $\mu = 1, \sigma^2 = 1.5$ D= {0, 9.81(0), 10.11(0.002), 40.00 (0.571), 40.03(0.004), 69.00(0.986), 70.15(1), 200.77(1)}. For $\mu = 2$ $\sigma^2 = 2.5,$ D= {0, 9.85(0), 10.11(0.002), 40.00 (0.571), 40.45 (0.578), 69.72(0.996), 72.73(1), 289.39 (1)}. For $\mu = 10$ $\sigma^2 = 10.5,$ D= {0, 9.52(0), 10.11(0.002), 40.00(0.571), 40.34(0.576), 69.00(0.986), 71.05(1), 238.14(1)}. For $\mu = 50$ $\sigma^2 = 55,$ D= {0, 9.81(0), 10.11(0.002), 9.52(0), 40.00(0.571), 40.03(0.004), 69.72(0.996), 71.05(1), 225.50(1)}. For $\mu = 100$ $\sigma^2 = 105,$ D= {0, 9.62(0), 10.11(0.002), 40.00(0.571), 40.16(0.574), 69.00(0.986), 70.15(1), 205.93(1)} respectively.

Table 4.4. Fuzzy set classification of different levels of overdispersion (PO) n=500, 1000.

n=500

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	0	9.81 (0)	10.24 (0.004)	40.00 (0.571)	40.03 (0.572)	69.72 (0.996)	71.05 (1)	246.05 (1)
0.05	0.1	0	9.81 (0)	10.24 (0.004)	39.92 (0.976)	40.41 (0.577)	69.72 (0.996)	70.15 (1)	151.60 (1)
0.5	0.55	0	9.85 (0)	10.16 (0.003)	40.00 (0.571)	40.16 (0.574)	69.72 (0.996)	70.51 (1)	245.45 (1)
1	1.5	0	9.81 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.004)	69.00 (0.986)	70.15 (1)	200.77 (1)
2	2.5	0	9.85 (0)	10.11 (0.002)	40.00 (0.571)	40.45 (0.578)	69.72 (0.996)	72.73 (1)	298.39 (1)
10	10.5	0	9.52 (0)	10.11 (0.002)	40.00 (0.571)	40.34 (0.576)	69.00 (0.986)	71.05 (1)	238.14 (1)
50	55	0	9.81 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.004)	69.72 (0.996)	71.05 (1)	225.50 (1)
100	105	0	9.62 (0)	10.11 (0.002)	40.00 (0.571)	40.16 (0.574)	69.00 (0.986)	70.15 (1)	205.93 (1)
		n=1000							
		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	0	9.85 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.72 (0.996)	70.15 (1)	189.13 (1)
0.05	0.1	0	9.85 (0)	10.11 (0.002)	40.00 (0.571)	40.16 (0.574)	69.00 (0.986)	70.15 (1)	233.18 (1)
0.5	0.55	0	9.85 (0)	10.01 (0.002)	40.00 (0.571)	40.20 (0.574)	69.72 (0.996)	71.05 (1)	255.71 (1)
1	1.5	0	9.81 (0)	10.11 (0.002)	40.00 (0.571)	40.21 (0.574)	69.72 (0.996)	71.05 (1)	205.93 (1)
2	2.5	0	9.62 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.72 (0.996)	71.05 (1)	287.20 (1)
10	10.5	0	9.85 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	68.35 (0.976)	71.05 (1)	245.45 (1)
50	55	0	9.81 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.72 (0.996)	70.15 (1)	258.20 (1)
100	105	0	9.81 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.00 (0.986)	71.05 (1)	287.20 (1)

Table 4.4 shows the results for the Fuzzy c- partition simulation for when n=1000, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01, \sigma^2 = 0.05,$,D={0, 9.85(0), 10.11(0.002), 40.00(0.571), 40.03(0.572), 69.72(0.996), 70.15(1), 189.13(1)}. For $\mu = 0.05, \sigma^2 = 0.1$, D= {0, 9.85(0), 10.01(0.002), 40.00(0.571), 40.16(0.574), 69.00(0.986), 70.15(1), 233.18(1)}. For $\mu = 0.5, \sigma^2 = 0.55$ D= {0, 9.85(0), 10.16(0.003), 40.00 (0.571), 40.20(0.574), 69.72(0.996), 71.05(1), 255.71(1)}. For $\mu = 1, \sigma^2 = 1.5$ D= {0, 9.81(0), 10.11(0.002), 40.00 (0.571), 40.21(0.574), 40.21(0.004), 71.05(1), 69.72(0.996), 70.05(1), 205.93(1)}. For $\mu = 2$ $\sigma^2 = 2.5$, D= {0, 9.62(0), 10.11(0.002), 40.00(0.571), 40.03(0.572), 69.72(0.996), 71.05(1), 287.20(1)}. For $\mu = 10$ $\sigma^2 = 10.5$, D= {0, 9.85(0), 10.11(0.002), 40.00(0.571), 40.03(0.572), 68.35(0.976), 71.05(1), 245.45(1)}. For $\mu = 50$ $\sigma^2 = 55$, D= {0, 9.81(0), 10.11(0.002), 40.00(0.571), 40.03(0.004), 69.72(0.996), 71.15(1), 258.20(1)}. For $\mu = 100$ $\sigma^2 = 105$, D= {0, 9.81(0), 10.11(0.002), 40.00(0.571), 40.03(0.0572), 69.00(0.986), 71.05(1), 287.20(1) } respectively.

Table 4.5 shows the results for the Fuzzy c- partition simulation for when n=5000, theminimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01, \sigma^2 = 0.05,$,D={0, 9.85(0), 10.11(0.002), 40.00(0.571), 40.03(0.572), 69.72(0.996), 70.15(1), 298.39 (1)}. For $\mu = 0.05, \sigma^2 = 0.1$, D= {0, 9.85(0), 10.01(0.002), 40.00(0.571), 40.03(0.572), 69.72(0.996), 70.15(1), 298.39 (1)}. For $\mu = 0.5, \sigma^2 = 0.55$ D= {0, 9.85(0), 10.16(0.003), 40.00 (0.571), 40.20(0.574), 69.72(0.996), 71.05(1), 255.71(1)}. For $\mu = 1, \sigma^2 = 1.5$ D= {0, 9.85(0), 10.07(0.002), 40.00 (0.571), 40.03(0.572), 69.87(0.998), 70.15(1), 451.61(1)}. For $\mu = 2$ $\sigma^2 = 2.5$, D= {0, 9.85(0), 10.11(0.002), 40.00 (0.571), 40.03(0.572), 69.87(0.998), 70.15(1), 375.00(1)}. For $\mu = 10$ $\sigma^2 = 10.5$, D= {0, 9.81(0), 10.11(0.002), 40.00(0.571), 40.03(0.572), 69.87(0.998), 70.15(1), 375.00 (1)}. For $\mu = 50$ $\sigma^2 = 55$, D= {0, 9.85(0), 10.01(0.002), 40.00(0.571), 40.11(0.0573), 69.87(0.998), 70.15(1), 241.03(1)}. For $\mu = 100$ $\sigma^2 = 105$, D= {0, 9.85(0), 10.01(0.002), 40.00(0.571), 40.03(0.0572), 69.87(0.998), 70.15(1), 411.54 (1)} respectively.

Table 4.5. Fuzzy set classification of different levels of overdispersion of (PO) n=5000.

		Not Severe		Moderately Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	0	9.85 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.72 (0.996)	70.15 (1)	298.39 (1)
0.05	0.1	0	9.85 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.72 (0.996)	70.15 (1)	298.39 (1)
0.5	0.55	0	9.85 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.72 (0.996)	70.15 (1)	311.11 (1)
1	1.5	0	9.85 (0)	10.07 (0.002)	40.00 (0.571)	40.03 (0.572)	69.87 (0.998)	70.15 (1)	451.61 (1)
2	2.5	0	9.85 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.87 (0.998)	70.15 (1)	307.14 (1)
10	10.5	0	9.81 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.87 (0.998)	70.15 (1)	375.00 (1)
50	55	0	9.85 (0)	10.01 (0.002)	40.00 (0.571)	40.11 (0.573)	69.87 (0.998)	70.15 (1)	241.03 (1)
100	105	0	9.85 (0)	10.01 (0.002)	40.00 (0.571)	40.03 (0.572)	69.87 (0.998)	70.15 (1)	411.54 (1)

NEGATIVE BINOMIAL MODEL

Table 4.6 Fuzzy set classification of different levels of overdispersion of NB n=20

		Not Severe		Moderately Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	-	-	-	-	57.63 (0.823)	68.95 (0.985)	71.87 (1)	86.41 (1)
0.05	0.1	-	-	34.80 (0.472)	-	46.29 (0.661)	68.37 (0.977)	71.99 (1)	83.23 (1)
0.5	0.55	-	-	-	-	56.27 (0.804)	67.61 (0.966)	71.62 (1)	90.54 (1)
1	1.5	-	-	-	-	51.35 (0.734)	67.41 (0.963)	70.42 (1)	89.82 (1)
2	2.5	-	-	-	-	68.76 (0.982)	69.67 (0.995)	71.46 (1)	91.04 (1)
10	10.5	-	-	-	-	57.56 (0.822)	68.32 (0.976)	71.16 (1)	91.08 (1)
50	55	-	-	-	-	41.49 (0.593)	69.09 (0.987)	70.62 (1)	89.94 (1)
100	105	-	-	-	-	53.89 (0.770)	58.35 (0.834)	70.28 (1)	91.54 (1)

Table 4.6-4.10 show the results for the Fuzzy c- partition simulation for NB; Table 4.6 shows the result when n=20, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01, \sigma^2 = 0.05,$,D={57.63(0.823), 68.95(0.985), 71.87(1), 86.41(1)}. For $\mu = 0.05, \sigma^2 = 0.1$, D= {34.80(0.472), 46.29(0.661), 68.37(0.977), 71.99(1), 83.23(1)}. For $\mu = 0.5, \sigma^2 = 0.55$ D= {56.27(0.804), 67.61(0.966), 71.62(1), 90.54(1)}. For $\mu = 1, \sigma^2 = 1.5$ D= {51.35(0.734), 67.41(0.963), 70.42(1), 89.82(1)}. For $\mu = 2$ $\sigma^2 = 2.5$, D= {68.76(0.982), 71.46(1), 91.04(1), 375.00(1)}. For $\mu = 10$ $\sigma^2 = 10.5$, D= {57.56(0.822), 68.32(0.976), 71.16(1), 91.08(1)}. For $\mu = 50$ $\sigma^2 = 55$, D= {41.49(0.593), 69.09(0.987), 70.62(1), 89.94(1)}. For $\mu = 100$ $\sigma^2 = 105$, D= {53.89(0.770), 58.35(0.834), 70.28(1), 91.54(1)} respectively.

Table 4.7 shows the results for the Fuzzy c- partition simulation for when n=30, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01, \sigma^2 = 0.05,$,D = {40.45(0.578), 69.97(0.999), 70.81(1),90.65(1)}. For $\mu = 0.05, \sigma^2 = 0.1$, D = {44.36(0.633), 69.98(0.999), 71.04(1),91.20(1)}. For $\mu = 0.5, \sigma^2 = 0.55$ D = {32.49(0.428), 34.89(0.474), 45.16 (0.645), 69.98(0.999), 70.36(1), 88.81(1)}. For $\mu = 1, \sigma^2 = 1.5$ D = {56.44(0.804), 66.27(0.947), 70.28(1), 88.76(1)}. For $\mu = 2$ $\sigma^2 = 2.5$, D = {55.70(0.976), 69.94(0.999), 70.34(1), 87.66(1)}. For $\mu = 10$ $\sigma^2 = 10.5$, D= {43.38(0.620), 69.11(0.987), 70.55(1), 90.20(1)}. For $\mu = 50$ $\sigma^2 = 55$, D = {45.36(0.648), 69.80(0.997), 72.17(1), 92.75(1)}. For $\mu = 100$ $\sigma^2 = 105$, D = {51.95(0.742), 68.67(0.981), 70.97(1), 90.39(1)} respectively.

Table 4.7 shows the results for the Fuzzy c- partition simulation for when n=50, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01, \sigma^2 = 0.05,$,D={49.21(0.703), 69.77(0.977), 70.33(1), 90.72(1)}. For $\mu = 0.05, \sigma^2 = 0.1$, D= {49.43(0.706), 69.36(0.990), 70.73(1), 90.26(1)}. For $\mu = 0.5, \sigma^2 = 0.55$ D= {41.84(0.598), 69.52(0.993), 70.42(1), 92.89(1)}. For $\mu = 1, \sigma^2 = 1.5$ D= {35.78(0.497), 49.70(0.710), 69.90(0.999), 92.06(1) }. For $\mu = 2$ $\sigma^2 = 2.5$, D= {59.02(0.843), 73.36(1), 90.88(1)}. For $\mu = 10$ $\sigma^2 = 10.5$, D= {48.28(0.690), 69.78(0.997), 70.54(1), 93.31(1)}. For $\mu = 50$ $\sigma^2 = 55$, D= {41.52(0.593), 69.13(0.988), 70.35(1), 93.95(1)}.

Table 4.7. Fuzzy set classification of different levels of overdispersion of NB n=30, 50.

n=30

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	-	-	-	-	49.21 (0.703)	69.77 (0.977)	70.33 (1)	90.72 (1)
0.05	0.1	-	-	-	-	49.43 (0.706)	69.36 (0.990)	70.73 (1)	90.26 (1)
0.5	0.55	-	-	-	-	41.84 (0.598)	69.52 (0.993)	70.42 (1)	92.89 (1)
1	1.5	-	-	35.78 (0.497)	-	49.70 (0.710)	69.90 (0.999)	70.18 (1)	92.06 (1)
2	2.5	-	-	-	-	59.02 (0.843)	-	73.36 (1)	90.88 (1)
10	10.5	-	-	-	-	48.28 (0.690)	69.78 (0.997)	70.54 (1)	93.31 (1)
50	55	-	-	-	-	41.52 (0.593)	69.13 (0.988)	70.35 (1)	93.95 (1)
100	105	-	-	-	-	45.91 (0.656)	69.80 (0.997)	70.16 (1)	88.65 (1)

n=50

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	-	-	32.68 (0.432)	-	46.59 (0.666)	69.87 (0.998)	70.01 (1)	94.28 (1)
0.05	0.1	-	-	34.93 (0.475)	39.68 (0.565)	40.23 (0.574)	69.45 (0.992)	70.28 (1)	92.23 (1)
0.5	0.55	7.86 (0)	-	29.26 (0.367)	-	44.12 (0.630)	69.84 (0.998)	70.42 (1)	92.89 (1)
1	1.5	-	-	11.99 (0.038)	-	48.46 (0.692)	69.28 (0.990)	70.18 (1)	92.06 (1)
2	2.5			17.39 (0.141)	-	41.97 (0.600)	69.91 (0.999)	71.90 (1)	90.88 (1)
10	10.5			31.93 (0.418)	-	47.64 (0.681)	69.78 (0.992)	70.54 (1)	93.31 (1)
50	55			-	-	42.13 (0.602)	69.39 (0.991)	70.35 (1)	93.95 (1)
100	105			-	-	45.91 (0.656)	69.86 (0.998)	70.16 (1)	88.65 (1)

For $\mu\ = 100\ \sigma^2\ =\ 105,\ D=$ {45.91(0.656), 69.80(0.997), 70.16(1), 88.65(1)} respectively.

Table 4.8 shows the results for the Fuzzy c- partition simulation for when n=100, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu\ = 0.01, \sigma^2\ =\ 0.05, ,D\ =$ {32.68(0.432), 46.59(0.666), 69.87(0.998) 70.01(1),94.28(1)}. For $\mu\ = 0.05, \sigma^2\ =\ 0.1,\ D=$ {34.93(0.475), 39.68(0.565), 40.23(0.574), 69.45(0.992), 70.28(1), 92.23(1)}. For $\mu\ = 0.5, \sigma^2\ =\ 0.55\ \ D\ =$ {7.86(0), 29.26(0.367), 44.12(0.630), 69.84(0.998), 70.42(1), 92.89(1)}. For $\mu\ = 1, \sigma^2\ =\ 1.5\ \ D\ =$ {11.99(0.038), 48.46(0.692), 69.28(0.990), 70.18(1), 92.06(1)}. For $\mu\ = 2\ \sigma^2\ =\ 2.5,\ D\ =$ {17.39(0.141), 41.97(0.600), 69.91(0.999), 71.90(1), 90.88(1)}. For $\mu\ = 10\ \sigma^2\ =\ 10.5,\ D\ =$ {31.93(0.418), 47.64(0.681), 69.78(0.992), 70.54(1), 93.31(1)}. For $\mu\ = 50\ \sigma^2\ =\ 55,\ D\ =$ {42.13(0.602), 69.39(0.991), 70.35(1), 93.95(1)}. For $\mu\ = 100\ \sigma^2\ =\ 105,\ D\ =$ {45.91(0.656), 69.86(0.998), 70.16(1), 88.65(1)} respectively.

Table 4.8 shows the results for the Fuzzy c- partition simulation for when n=200, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu\ = 0.01, \sigma^2\ =\ 0.05, ,D\ =$ {39.60(0.564), 45.80(0.654), 69.99(0.999), 70.50(1), 92.55(1)}. For $\mu\ = 0.05, \sigma^2\ =\ 0.1,\ D\ =$ {16.53(0.124), 35.90(0.493), 40.20(0.574), 69.93(0.999), 70.02(1), 93.51(1)}. For $\mu\ = 0.5, \sigma^2\ =\ 0.55\ \ D\ =$ {20.98(0.209), 44.33(0.633), 69.45(0.992), 70.05(1) 92.25(1)}. For $\mu\ = 1, \sigma^2\ =\ 1.5\ \ D\ =$ {38.71(0.547), 41.44(0.592), 69.96(0.999), 70.02(1), 90.32(1)}. For $\mu\ = 2\ \sigma^2\ =\ 2.5,\ D\ =$ {8.50(0), 36.90(0.512), 43.74(0.625), 69.92(0.999), 70.14(1), 92.52(1)}. For $\mu\ = 10\ \sigma^2\ =\ 10.5,\ D=$ {23.07(0.249), 34.48(0.466), 43.14(0.616), 69.97(0.999), 70.33(1), 94.40(1)}. For $\mu\ = 50\ \sigma^2\ =\ 55,\ D\ =$ {41.04(0.586), 69.97(0.999), 70.12(1), 91.69(1)}. For $\mu\ = 100\ \sigma^2\ =\ 105,\ D\ =$ {45.91(0.656), 69.57(0.994), 70.06(1), 92.06(1)} respectively.

Table 4.9 shows the results for the Fuzzy c- partition simulation for when n=300, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu\ = 0.01, \sigma^2\ =\ 0.05, ,D\ =$ {18.63(0.164), 36.31(0.501), 42.48(0.607), 39.60(0.564), 69.91(0,699), 70.08(1), 95.27(1)}. For $\mu\ = 0.05, \sigma^2\ =\ 0.1,\ D\ =$ {32.41(0.427), 33.86(0.454), 41.26(0.598), 69.85(0.988), 70.05(1), 93.57(1)}. For $\mu\ = 0.5, \sigma^2\ =\ 0.55\ \ D$ = {22.88(0.245), 25.12(0.288), 47.59(0.680),

Table 4.8. Fuzzy set classification of different levels of overdispersion of NB when n=100,200

n=100

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	-	-	32.68 (0.432)	-	46.59 (0.666)	69.87 (0.998)	70.01 (1)	94.28 (1)
0.05	0.1	-	-	34.93 (0.475)	39.68 (0.565)	40.23 (0.574)	69.45 (0.992)	70.28 (1)	92.23 (1)
0.5	0.55	7.86 (0)	-	29.26 (0.367)	-	44.12 (0.630)	69.84 (0.998)	70.42 (1)	92.89 (1)
1	1.5	-	-	11.99 (0.038)	-	48.46 (0.692)	69.28 (0.990)	70.18 (1)	92.06 (1)
2	2.5			17.39 (0.141)	-	41.97 (0.600)	69.91 (0.999)	71.90 (1)	90.88 (1)
10	10.5			31.93 (0.418)	-	47.64 (0.681)	69.78 (0.992)	70.54 (1)	93.31 (1)
50	55			-	-	42.13 (0.602)	69.39 (0.991)	70.35 (1)	93.95 (1)
100	105			-	-	45.91 (0.656)	69.86 (0.998)	70.16 (1)	88.65 (1)

n=200

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	-	-	39.60 (0.564)	-	45.80 (0.654)	69.99 (0.999)	70.50 (1)	92.55 (1)
0.05	0.1	-	-	16.53 (0.124)	35.90 (0.493)	40.20 (0.574)	69.93 (0.999)	70.02 (1)	93.51 (1)
0.5	0.55	-	-	20.98 (0.209)	-	44.33 (0.633)	69.45 (0.992)	70.05 (1)	92.25 (1)
1	1.5	-	-	38.71 (0.547)	-	41.44 (0.592)	69.96 (0.999)	70.02 (1)	90.32 (1)
2	2.5	8.50 (0)	-	36.90 (0.512)	-	43.74 (0.625)	69.92 (0.999)	70.14 (1)	92.52 (1)
10	10.5	-	-	23.07 (0.249)	34.48 (0.466)	43.14 (0.616)	69.97 (0.999)	70.33 (1)	94.40 (1)
50	55	-	-	-	-	41.04 (0.586)	69.97 (0.999)	70.12 (1)	91.69 (1)
100	105	-	-	-	-	45.19 (0.646)	69.57 (0.994)	70.06 (1)	92.06 (1)

69.55(0.994), 70.27(1), 94.22(1)}. For $\mu = 1, \sigma^2 = 1.5$ D = {18.36(0.159), 19.50(0.181), 42.42(0.915), 69.94(0.999), 70.00(1), 92.91(1)}. For $\mu = 2$ $\sigma^2 = 2.5$, D = {25.77(0.300), 34.39(0.465), 44.28(0.837), 69.79(0.994), 70.00(1), 92.59(1)}. For $\mu = 10$ $\sigma^2 = 10.5$, D = {21.01(0.210), 35.30(0.482), 43.28(0.618), 23.07(0.249), 69.97(0.999), 70.02(1), 95.88(1)}.For $\mu = 50$ $\sigma^2 = 55$, D = {24.72(0.280), 26.92(0.322), 42.33(0.605), 69.97(0.999), 70.10(1), 92.86(1)}. For $\mu = 100$ $\sigma^2 = 105$, D = {35.96(0.494), 40.11(0.573), 69.99(0.999), 70.51(1), 93.50(1)} respectively.

Table 4.9 shows the results for the Fuzzy c- partition simulation for when n=500, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01, \sigma^2 = 0.05,$,D={23.46(0.256), 36.13(0.498), 43.13(0.616), 69.95(0.999), 70.09(1), 93.25(1)}. For $\mu = 0.05, \sigma^2 = 0.1$, D= {20.83(0.206), 39.51(0.562), 41.36(0.591), 69.86(0.998), 70.10(1), 95.96(1)}. For $\mu = 0.5, \sigma^2 = 0.55$ D= {30.23(0.385), 37.36(0.521). 44.12(0.630), 69.97(0.999), 70.00(1), 93.29(1)}. For $\mu = 1, \sigma^2 = 1.5$ D= {30.00(0.381), 39.21(0.556), 40.35(0.666), 69.92(0.999), 70.00(1), 95.36(1)}. For $\mu = 2$ $\sigma^2 = 2.5$, D= {15.99(0.114), 38.85(0.549), 44.27(0.632), 69.75(0.996), 70.12(1), 95.11(1)}. For $\mu = 10$ $\sigma^2 = 10.5$, D= {22.97(0.247), 38.87(0.550), 41.84(0.598), 69.84(0.998), 70.08(1), 92.69(1)}. For $\mu = 50$ $\sigma^2 = 55$, D= {21.17(0.213), 39.78(0.567), 41.64(0.595). 69.93(0.999), 70.02(1), 94.39(1)}. For $\mu = 100$ $\sigma^2 = 105$, D= {13.64(0.069), 38.71(0.547), 40.48(0.578), 69.91(0,999), 70.16(1), 93.22(1) respectively.

Table 4.10. shows the results for the Fuzzy c- partition simulation for when n=1000, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01$, $\sigma^2 = 0.05$, D={10.47(0.009), 39.36(0.559), 41.29(0.590), 69.97(0.999), 70.10(1), 96.18(1)}. For $\mu = 0.05, \sigma^2 = 0.1$, D= {17.04(0.134), 39.98 (0.571), 40.26(0.575), 69.99(0.999), 70.12(1), 95.03(1)}. For $\mu = 0.5, \sigma^2 = 0.55$ D= {6.12(0), 8.72(0), 16.79(0.134), 39.04(0.553), 41.36(0.574), 69.93(0.999), 70.05(1), 95.05(1)}. For $\mu = 1, \sigma^2 = 1.5$ D= {5.41(0), 36.36(0.559), 41.36(0.591), 69.92(0.999), 72.73(1), 145.65(1)}. For $\mu = 2$ $\sigma^2 = 2.5$, D= {0, 9.62(0), 10.16(0.003), 40.00(0.571), 40.03(0.572), 68.40(0.977), 72.73(1), 249.88(1)}. For $\mu = 10$ $\sigma^2 = 10.5$, D= {0.05(0), 9.96(0), 10.00(0.002),

Table 4.9. Fuzzy set classification of different levels of overdispersion of NB n=300, 500

n=300

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	-	-	18.63 (0.164)	36.31 (0.501)	42.48 (0.607).	69.91 (0,699)	70.08 (1)	95.27 (1)
0.05	0.1	-	-	32.41 (0.427)	33.86 (0.454)	41.26 (0.598)	69.85 (0.988)	70.05(1)	93.57 (1)
0.5	0.55	-	-	22.88 (0.245)	25.12 (0.288)	47.59 (0.680)	69.55 (0.994)	70.27 (1)	94.22 (1)
1	1.5	-	-	18.36 (0.159)	19.50 (0.181)	42.42 (0.915)	69.94 (0.999)	70.00 (1)	92.91 (1)
2	2.5	-	-	25.77 (0.300)	34.39 (0.465)	44.28 (0.837)	69.79 (0.994)	70.00 (1)	92.59 (1)
10	10.5	-	-	21.01 (0.210)	35.30 (0.482)	43.28 (0.618)	69.97 (0.999)	70.02 (1)	95.88 (1)
50	55	-	-	24.72 (0.280)	26.92 (0.322)	42.33 (0.605)	69.97 (0.999)	70.10 (1)	92.86 (1)
100	105	-	-	35.96 (0.494)	-	40.11 (0.573)	69.99 (0.999)	70.51 (1)	93.50 (1)

n=500

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	-	-	23.46 (0.256)	36.13 (0.498)	43.13 (0.616)	69.95 (0.999)	70.09 (1)	93.25 (1)
0.05	0.1	-	-	20.83 (0.206)	39.51 (0.562)	41.36 (0.591)	69.86 (0.998)	70.10 (1)	95.96 (1)
0.5	0.55	-	-	30.23 (0.385)	37.36 (0.521)	44.12 (0.630)	69.97 (0.999)	70.00 (1)	93.29 (1)
1	1.5	-	-	30.00 (0.381)	39.21 (0.556)	40.35 (0.666)	69.92 (0.999)	70.00 (1)	95.36 (1)
2	2.5	6.37(0)	-	15.99 (0.114)	38.85 (0.549)	44.27 (0.632)	69.75 (0.996)	70.12 (1)	95.11 (1)
10	10.5	-	-	22.97 (0.247)	38.87 (0.550)	41.84 (0.598)	69.84 (0.998)	70.08 (1)	92.69 (1)
50	55	0.85(0)	-	21.17 (0.213)	39.78 (0.567)	41.64 (0.595)	69.93 (0.999)	70.02 (1)	94.39 (1)
100	105	-	-	13.64 (0.069)	38.71 (0.547)	40.48 (0.578)	69.91 (0,999)	70.16 (1)	93.22 (1)

Table 4.10. Fuzzy set classification of different levels of overdispersion of NB

n=1000, 5000

n=1000

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	-	-	10.47 (0.009)	39.36 (0.559)	41.29 (0.590)	69.97 (0.999)	70.10 (1)	96.18 (1)
0.05	0.1	-		17.04 (0.134)	39.98 (0.571	40.26 (0.575)	69.99 (0.999)	70.12 (1)	95.03 (1)
0.5	0.55	6.12 (0)	8.72 (0)	16.79 (0.134)	39.04 (0.553)	41.36 (0.574)	69.93 (0.999)	70.05 (1)	95.05 (1)
1	1.5	5.41 (0)	-	20.96 (0.209)	36.36 (0.559)	41.36 (0.591	69.92 (0.999)	72.73 (1)	145.65 (1)
2	2.5	0 (0)	9.62 (0)	10.16 (0.003)	40.00 (0.571)	40.03 (0.572)	68.40 (0.977)	72.73 (1)	249.88 (1)
10	10.5	0.05 (0)	9.96 (0)	10.00 (0.002)	39.91 (0.570)	40.20 (0.574)	69.53 (0.993)	70.33 (1)	249.88 (1)
50	55	0.01 (0)	9.99 (0)	10.01 (0.002)	39.99 (0.577)	40.07 (0.572)	69.24 (0.989)	70.57 (1)	165.03 (1)
100	105	0.09 (0)	9.88 (0)	10.05 (0.001)	39.97 (0.571)	40.03 (0.572)	69.78 (0.997)	70.12 (1)	78.16 (1)

n=5000

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	0.65 (0)	7.54 (0)	11.33 (0.025)	39.87 (0.569)	40.23 (0.782)	69.99 (0.984)	70.00 (1)	97.86 (1)
0.05	0.1	4.32 (0)	9.85 (0)	13.07 (0.058)	39.78 (0.567)	40.23 (0.883)	69.99 (0.984)	70.00 (1)	96.42 (1)
0.5	0.55	0	9.52 (0)	10.94 (0.018)	39.90 (0.570)	40.00 (0.571)	69.91 (0.999)	70.10 (1)	111.11 (1)
1	1.5	0	9.52 (0)	10.16 (0.003)	39.97 (0.571)	40.00 (0.571)	69.88 (0.998)	70.05 (1)	171.43 (1)
2	2.5	0	9.85 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.73 (0.996)	70.26 (1)	161.59 (1)
10	10.5	0	10.0 (0)	10.02 (0.002)	40.00 (0.571)	40.03 (0.572)	69.92 (0.999)	70.01 (1)	139.36 (1)
50	55	0.04 (0)	9.98 (0)	10.01 (0.002)	39.99 (0.577)	40.00 (0.571)	69.96 (0.999)	70.00 (1)	169.40 (1)
100	105	0.07 (0)	9.98 (0)	10.01 (0.002)	39.98 (0.571)	40.01 (0.572)	69.95 (0.999)	70.01 (1)	169.88 (1)

39.91(0.570), 40.20(0.574), 69.53(0.993), 70.33(1), 249.88(1)}. For $\mu = 50\ \sigma^2 = 55,\ D$= {0.01(0), 9.99(0), 10.01(0.002), 39.99(0.577), 40.07(0.572), 69.24(0.989), 70.57(1), 165.03(1)}. For $\mu = 100\ \sigma^2 = 105,\ D$= {0.09(0), 9.88(0), 10.05(0.001), 39.97(0.571), 40.03(0.572), 69.78(0.997), 70.12(1), 78.16(1)} respectively.

Table 4.10 shows the results for the Fuzzy c- partition simulation for when n=5000, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01,\ \sigma^2 = 0.05, D$={0.65(0), 7.54(0), 11.33(0.025), 39.87(0.569), 40.23(0.782), 69.99(0.984), 70.00(1), 97.86(1)}. For $\mu = 0.05, \sigma^2 = 0.1,\ D$= {4.32(0), 9.85(0), 13.07(0.058), 39.78(0.567), 40.23(0.883), 69.99(0.984), 70.00(1), 96.42(1)}. For $\mu = 0.5, \sigma^2 = 0.55\ D$= {0, 9.52(0), 10.94(0.018), 39.90(0.570), 40.00(0.571), 69.91(0.999), 70.10(1), 111.11(1)}. For $\mu = 1, \sigma^2 = 1.5\ D$= {0, 9.52(0), 10.16(0.003), 39.97(0.571), 40.00(0.571), 69.88(0.998), 70.05(1), 171.43(1))}. For $\mu = 2\ \sigma^2 = 2.5,\ D$= {0, 9.85(0), 10.11(0.002), 40.00(0.571), 40.03(0.572), 69.73(0.996), 70.26(1), 161.59(1)}. For $\mu = 10\ \sigma^2 = 10.5,\ D$= {0, 10.00(0.002), 10.02(0.002), 40.00(0.571), 40.03(0.572), 69.92(0.999), 70.01(1), 139.36(1)}. For $\mu = 50\ \sigma^2 = 55,\ D$= {0.04(0), 9.98(0), 10.01(0.002), 39.99(0.577), 40.00(0.571), 69.96(0.999), 70.00(1), 169.40(1)} For $\mu = 100\ \sigma^2 = 105,\ D$= {0.07(0), 9.98(0), 10.01(0.002), 39.98(0.571), 40.01(0.572), 69.95(0.999), 70.01(1), 169.88(1)} respectively.

Table 4.11-4.15 show the results for the Fuzzy c- partition simulation for CP; Table 4.11 shows the result when n=20, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01, \sigma^2 = 0.05,\ D$ = {11.20(0.233, 37.25(0.591), 47.68(0.681), 39.87(0.569), 63.91(0.913), 130.15(1)}. For $\mu = 0.05, \sigma^2 = 0.1,\ D$= {0.48(0), 7.55(0), 14.77(0.091), 39.94(0.570), 45.12(0.645), 58.50(0.836)}. For $\mu = 0.5, \sigma^2 = 0.55\ D$= {0.43(0), 4.32(0), 10.08(0.002), 38.27(0.538), 42.42(0.606), 44.72(0.639)}. For $\mu = 1, \sigma^2 = 1.5\ D$= {1.22(0), 9.64(0), 19.85(0.188), 34.48(0.466), 41.52(0.593), 44.99(0.643), 99.24(1)}. For $\mu = 2\ \sigma^2 = 2.5,\ D$= {17.53(0.143), 37.93(0.532), 40.23(0.575), 69.30(0.990), 70.15(1), 86.22(1)}. For $\mu = 10\ \sigma^2 = 10.5,\ D$= {3.61 0), 9.13(0), 12.94(0.056), 38.69(0.546), 42.03(0.600)}. For $\mu = 50\ \sigma^2 = 55,\ D$= {2.12(0), 4.21(0), 0.04(0), 10.08(0.002), 38.25(0.538), 40.16(0.574), 54.22(0.775)}. For $\mu = 100\ \sigma^2 = 105,$

Com-Poisson Model

Table 4.11. Fuzzy set classification of different levels of overdispersion of CP.

n=20, 30

n=20

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	-	-	11.20 (0.233)	37.25 (0.517)	47.68 (0.681)	63.91 (0.913)	130.15 (1)	-
0.05	0.1	0.48 (0)	7.55 (0)	14.77 (0.091)	39.94 (0.570)	45.12 (0.645)	58.50 (0.836)	-	-
0.5	0.55	0.43 (0)	4.32 (0)	10.08 (0.002)	38.27 (0.538)	42.42 (0.606)	44.72 (0.639)	-	-
1	1.5	1.22 (0)	9.64 (0)	19.85 (0.188)	34.48 (0.466)	41.52 (0.593)	44.99 (0.643)	99.24 (1)	-
2	2.5	-	-	17.53 (0.143)	37.93 (0.532)	40.23 (0.575)	69.30 (0.990)	70.15 (1)	86.22 (1)
10	10.5	3.619 (0)	9.13 (0)	12.94 (0.056)	38.69 (0.546)	42.03 (0.600)	-	-	-
50	55	2.12 (0)	4.21 (0)	10.08 (0.002)	38.25 (0.538)	40.16 (0.574)	54.22 (0.775)	-	-
100	105	0.09 (0)	8.70 (0)	13.50 (0.067)	39.44 (0.561)	40.38 (0.577)	61.11 (0.873)	-	-
		n=30							
		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	2.65 (0)	9.09 (0)	13.90 (0.074)	38.44 (0.542)	40.97 (0.585)	52.48 (0.613)	91.91 (1)	-
0.05	0.1	0.49 (0)	9.27 (0)	10.75 (0.014)	38.70 (0.547)	40.44 (0.578)	55.65 (0.974)	-	-
0.5	0.55	1.69 (0)	9.16 (0)	10.26 (0.005)	38.57 (0.544)	41.22 (0.589)	48.11 (0.687)	-	-
1	1.5	2.14 (0)	9.94 (0)	11.00 (0.019)	38.95 (0.551)	44.61 (0.637)	53.20 (0.760)	91.80 (1)	-
2	2.5	0.71 (0)	9.18 (0)	10.21 (0.004)	39.33 (0.559)	44.19 (0.631)	51.81 (0.740)	-	-
10	10.5	1.63 (0)	8.79 (0)	10.55 (0.010)	38.62 (0.545)	43.20 (0.617)	67.20 (0.960)	-	-
50	55	2.76 (0)	9.65 (0)	10.59 (0.011)	38.40 (0.541)	44.86 (0.641)	46.23 (0.660)	-	-
100	105	0.46 (0)	5.54 (0)	13.38 (0.064)	39.42 (0.560)	40.75 (0.582)	55.48 (0.793)	74.05 (1)	-

D= {0.09(0), 8.70(0), 13.50(0.067), 39.44(0.561), 40.38(0.577), 61.11(0.873)} respectively.

Table 4.11 shows the results for the Fuzzy c- partition simulation for when n=30, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01, \sigma^2 = 0.05, ,D$ = {2.65(0), 9.09(0), 13.90(0.074), 38.44(0.542), 40.97(0.585), 52.48(0.613), 91.91(1)}. For $\mu = 0.05, \sigma^2 = 0.1,\ D$= {0.49(0), 9.27(0), 10.75(0.014), 38.70(0.547), 40.44(0.578), 55.65(0.974)}. For $\mu = 0.5, \sigma^2 = 0.55\ D$= {1.69(0), 9.16(0), 10.26(0.005), 38.57(0.544), 41.22(0.589), 48.11(0.687)}. For $\mu = 1, \sigma^2 = 1.5\ D$= {2.14(0), 9.94(0), 11.00(0.019), 38.95(0.551), 44.61(0.637), 53.20(0.760), 91.80(1)}. For $\mu = 2\ \sigma^2 = 2.5,\ D$= {0.71(0), 9.18(0), 10.21(0.004), 39.33(0.559), 44.19(0.631), 51.81(0.740)}. For $\mu = 10\ \sigma^2 = 10.5,\ D$= {1.63(0), 8.79(0), 10.55(0.010), 38.62(0.545), 43.20(0.617), 67.20(0.960)}. For $\mu = 50\ \sigma^2 = 55,\ D$= {2.76(0), 9.65(0), 10.59(0.011), 38.40(0.541), 44.86(0.641), 46.23(0.660)}. For $\mu = 100\ \sigma^2 = 105,\ D$= {0.46(0), 5.54(0), 0.09(0), 13.38(0.064), 39.42(0.560), 40.75(0.582), 55.48(0.793), 74.05(1)} respectively.

Table 4.12 shows the results for the Fuzzy c- partition simulation for when n=50, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01,\ \sigma^2 = 0.05, D$ = {0.04(0), 9.95(0), 10.05(0.001), 39.39(0.560), 40.41(0.577)}. For $\mu = 0.05, \sigma^2 = 0.1,\ D$ = {0.01(0), 9.53(0), 39.88(0.569), 39.88(0.569), 42.17(0.602), 46.18(0.660), 72.94(1)}. For $\mu = 0.5, \sigma^2 = 0.55\ D$ = {4.68(0), 9.99(0), 12.73(0.052), 39.35(0.559), 40.80(0.583), 51.55(0.736)}. For $\mu = 1, \sigma^2 = 1.5\ D$ = {1.34(0), 8.27(0), 12.92(0.056), 38.94(0.551), 41.11(0.587), 48.26(0.689)}. For $\mu = 2\ \sigma^2 = 2.5,\ D$= {0.23(0), 9.17(0), 10.01(0.000), 39.75(0.567), 41.55(0.594), 52.05(0.744)}. For $\mu = 10\ \sigma^2 = 10.5,\ D$ = {0.73(0), 9.36(0), 12.09(0.040), 39.88(0.569), 40.20(0.574), 52.51(0.759)}. For $\mu = 50\ \sigma^2 = 55,\ D$ = {1.64(0), 8.02(0), 11.68(0.032), 36.31(0.501), 41.62(0.595), 53.50(0.764)}. For $\mu = 100\ \sigma^2 = 105,\ D$ = {3.34(0), 8.68(0), 0.46(0), 10.04(0.001), 39.46(0.561), 45.32(0.647)} respectively.

Table 4.12 shows the results for the Fuzzy c- partition simulation for when n=100, the minimum and maximum values for not severe, moderately severe, severe and very

Table 4.12. Fuzzy set classification of different levels of overdispersion of CP, n=50, 100

n=50

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	0.04 (0)	9.95 (0)	10.05 (0.001)	39.39 (0.560)	40.41 (0.577)	-	-	-
0.05	0.1	0.01 (0)	9.53 (0)	11.34 (0.026)	39.88 (0.569)	42.17 (0.602)	46.18 (0.660)	72.94 (1)	-
0.5	0.55	4.68 (0)	9.99 (0)	12.73 (0.052)	39.35 (0.559)	40.80 (0.583)	51.55 (0.736)	-	-
1	1.5	1.34 (0)	8.27 (0)	12.92 (0.056)	38.94 (0.551)	41.11 (0.587)	48.26 (0.689)	-	-
2	2.5	0.23 (0)	9.17 (0)	10.01 (0.000)	39.75 (0.567)	41.55 (0.594)	52.05 (0.744)	-	-
10	10.5	0.73 (0)	9.36 (0)	12.09 (0.040)	39.88 (0.569)	40.20 (0.574)	52.51 (0.759)	-	-
50	55	1.64 (0)	8.02 (0)	11.68 (0.032)	36.31 (0.501)	41.62 (0.595)	53.50 (0.764)	-	-
100	105	3.34 (0)	8.68 (0)	10.04 (0.001)	39.46 (0.561)	45.32 (0.647)	-	-	-
		n=100							
		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	0.20 (0)	9.42 (0)	10.40 (0.008)	39.64 (0.565)	40.85 (0.584)	-	-	-
0.05	0.1	0.50 (0)	9.42 (0)	10.39 (0.008)	36.53 (0.505)	41.53 (0.593)	46.23 (0.660)	-	-
0.5	0.55	0.20 (0)	9.30 (0)	10.17 (0.003)	39.74 (0.567)	40.19 (0.574)	43.49 (0.621)	-	-
1	1.5	1.30 (0)	9.20 (0)	10.66 (0.012)	38.84 (0.549)	40.90 (0.584)	47.68 (0.681)	-	-
2	2.5	0.14 (0)	9.13 (0)	11.19 (0.023)	39.56 (0.563)	49.21 (0.703)	-	-	-
10	10.5	0.52 (0)	9.83 (0)	10.66 (0.012)	39.18 (0.556)	45.89 (0.656)	-	-	-
50	55	0.22 (0)	9.85 (0)	10.00 (0.000)	39.24 (0.557)	40.24 (0.575)	-	-	-
100	105	0.67 (0)	7.45 (0)	10.00 (0.000)	38.75 (0.548)	40.96 (0.585)	47.74 (0.682)		-

severe respectively, for $\mu = 0.01, \sigma^2 = 0.05$, D = {0.20(0), 9.42(0), 10.40(0.008), 39.64(0.565), 40.85(0.584)}.

For $\mu = 0.05, \sigma^2 = 0.1$, D = {0.50(0), 9.42(0), 10.39(0.008), 36.53(0.505), 46.23(0.660)}. For $\mu = 0.5, \sigma^2 = 0.55$ D = {0.20(0), 9.30(0), 10.17(0.003), 39.74(0.567), 40.19(0.574), 43.49(0.621)}. For $\mu = 1, \sigma^2 = 1.5$ D= {1.30(0), 9.20(0), 10.66(0.012), 38.84(0.549), 40.90(0.584), 47.68(0.681)}. For $\mu = 2$ $\sigma^2 = 2.5$, D = {00.14(0), 9.13(0), 11.19(0.023), 39.56(0.563), 49.21(0.703)}. For $\mu = 10$ $\sigma^2 = 10.5$, D = {0.52(0), 9.83(0), 10.66(0.012), 39.18(0.556), 45.89(0.656)}. For $\mu = 50$ $\sigma^2 = 55$, D = {0.22(0), 9.85(0), 10.00(0.000), 39.24(0.557), 40.24(0.575)}. For $\mu = 100$ $\sigma^2 = 105$, D = {0.67(0), 7.45(0), 10.00(0.000), 38.75(0.548), 40.96(0.585), 47.74(0.682)} respectively.

Table 4.13 shows the results for the Fuzzy c- partition simulation for when n=200, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01$, $\sigma^2 = 0.05$, D = {1.83(0), 9.83(0), 10.28(0.006), 39.81(0.568, 42.78(0.611), }. For $\mu = 0.05, \sigma^2 = 0.1$, D = {0.07(0), 9.90(0), 10.280.006), 37.07(0.516), 41.61(0.594)}. For $\mu = 0.5, \sigma^2 = 0.55$ D = {0.03(0), 9.84(0), 10.05(0.001), 39.91(0.570)}. For $\mu = 1$, $\sigma^2 = 1.5$ D = {1.67(0), 9.99(0), 10.06(0.001), 38.46(0.542)}. For $\mu = 2$ $\sigma^2 = 2.5$, D = {0.98(0), 9.98(0), 10.12(0.002), 39.97(0.571)}. For $\mu = 10$ $\sigma^2 = 10.5$, D = {0.20(0), 9.92(0), 10.28(0.005), 39.18(0.500), 41.75(0.596)}. For $\mu = 50$ $\sigma^2 = 55$, D = {0.22(0), 9.84(0), 10.38(0.007), 37.46(0.523), 41.10(0.587)}. For $\mu = 100$ $\sigma^2 = 105$, D = {1.72(0), 9.89(0), 10.57(0.011), 39.01(0.552), 42.56(0.608)} respectively.

Table 4.13 shows the results for the Fuzzy c- partition simulation for when n=300, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01$, $\sigma^2 = 0.05$, D = {2.23(0), 9.41(0), 10.32(0.006), 37.90(0.531), 40.47(0.578)}. For $\mu = 0.05$, $\sigma^2 = 0.1$, D = {0.43(0), 9.41(0), 10.05(0.001), 10.58(0.011), 37.59(0.526)}. For $\mu = 0.5$, $\sigma^2 = 0.55$ D = {3.98(0), 91(0), 0.03(0), 9.84(0), 38.95(0.511)}. For $\mu = 1, \sigma^2 = 1.5$ D = {0.92(0), 9.42(0), 10.15(0.003), 36.41(0.503)}. For $\mu = 2$ $\sigma^2 = 2.5$, D = {2.68(0), 9.71(0), 10.12(0.002), 36.22(0.499)}. For $\mu = 10$ $\sigma^2 = 10.5$, D = {1.53(0), 9.84(0), 10.57(0.011), 36.56(0.506)}. For $\mu = 50$ $\sigma^2 = 55$, D = {5.66(0), 9.87(0),

Table 4.13. Fuzzy set classification of different levels of overdispersion of CP n=200, 300

n=200

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	1.83 (0)	9.83 (0)	10.28 (0.006)	39.81 (0.568)	42.78 (0.611)	-	-	-
0.05	0.1	0.07 (0)	9.90 (0)	10.28 (0.006)	37.07(0. 516	41.61 (0.594)	-	-	-
0.5	0.55	0.03 (0)	9.84 (0)	10.05 (0.001)	39.91 (0.570)	-	-	-	-
1	1.5	1.67 (0)	9.99 (0)	10.06 (0.001)	38.46 (0.542)	-	-	-	-
2	2.5	0.98 (0)	9.98 (0)	10.12 (0.002)	39.97 (0.571)	-	-	-	-
10	10.5	0.20 (0)	9.92 (0)	10.28 (0.005)	39.18 (0.500)	41.75 (0.596)	-	-	-
50	55	0.22 (0)	9.84 (0)	10.38 (0.007)	37.46 (0.523)	41.10 (0.587)	-	-	-
100	105	1.72 (0)	9.89 (0)	10.57 (0.011)	39.01 (0.552)	42.56 (0.608)	-	-	-

n=300

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	2.23 (0)	9.41 (0)	10.32 (0.006)	37.90 (0.531)	40.47 (0.578)	-	-	-
0.05	0.1	0.43 (0)	9.41 (0)	10.58 (0.011)	37.59 (0.526)	-	-	-	-
0.5	0.55	3.98 (0)	9.91 (0)	10.05 (0.001)	38.95 (0.511)	-	-	-	-
1	1.5	0.92 (0)	9.42 (0)	10.15 (0.003)	36.41 (0.503)	-	-	-	-
2	2.5	2.68 (0)	9.71 (0)	10.12 (0.002)	36.22 (0.499)	-	-	-	-
10	10.5	1.53 (0)	9.84 (0)	10.57 (0.011)	36.56 (0.506)	-	-	-	-
50	55	5.66 (0)	9.87 (0)	10.18 (0.003)	37.08 (0.516)	-	-	-	-
100	105	3.38 (0)	9.81(0)	10.03 (0.001)	37.41 (0.522)	-	-	-	-

Table 4.14. Fuzzy set classification of different levels of overdispersion of CP

n=500, B=1000

n=500

		Not Severe		Moderately Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	0.62 (0)	9.96 (0)	10.37 (0.007)	34.97 (0.476)	-	-	-	-
0.05	0.1	6.50 (0)	9.96 (0)	11.01 (0.019)	36.93 (0.513)	-	-	-	-
0.5	0.55	7.39 (0)	9.95 (0)	10.01 (0.000)	34.73 (0.471)	-	-	-	-
1	1.5	3.40 (0)	6.19 (0)	10.15 (0.003)	33.10 (0..440)	-	-	-	-
2	2.5	8.21 (0)	9.97 (0)	10.07 (0.001)	35.57 (0.487)	-	-	-	-
10	10.5	6.99 (0)	9.93 (0)	10.63 (0.012)	37.29 (0.520)	-	-	-	-
50	55	6.57 (0)	9.88 (0)	10.49 (0.009)	34.89 (0.474)	-	-	-	-
100	105	6.36 (0)	9.88 (0)	10.06 (0.001)	38.46 (0.542)	-	-	-	-

n=1000

		Not Severe		Moderately Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	7.48 (0)	10.60 (0)	34.20 (0.011)	37.48 (0.461)	-	-	-	-
0.05	0.1	-	-	11.10 (0,021)	32.13 (0.422)	-	-	-	-
0.5	0.55	8.11 (0)	-	10.82 (0.016)	32.14 (0.422)	-	-	-	-
1	1.5	8.23 (0)	-	11.73 (0.033)	32.64 (0.431)	-	-	-	-
2	2.5	-	-	10.93 (0.018)	32.34 (0.426)	-	-	-	-
10	10.5	-	-	10.22 (0.004)	31.14 (0.403)	-	-	-	-
50	55	7.74 (0)	-	10.62 (0.012)	33.09 (0.440)	-	-	-	-
100	105	8.76 (0)	-	10.94 (0.018)	31.83 (0.416)	-	-	-	-

Table 4.15. Fuzzy set classification of different levels of overdispersion of CP

n=5000

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	-	-	16.07 (0.116)	28.46 (0.352)	-	-	-	-
0.05	0.1	-	-	16.53 (0.124)	27.49 (0.333)	-	-	-	-
0.5	0.55	-	-	16.07 (0.115)	27.77 (0.388)	-	-	-	-
1	1.5	-	-	15.64 (0.107)	27.68 (0.337)	-	-	-	-
2	2.5	-	-	16.06 (0.115)	28.46 (0.352)	-	-	-	-
10	10.5	-	-	16.53 (0.124)	27.49 (0.333)	-	-	-	-
50	55	-	-	16.07 (0.116)	27.77 (0.338)	-	-	-	-
100	105	-	-	15.64 (0.107)	27.68 (0.337)	-	-	-	-

10.18(0.003), 37.08(0.516)}. For $\mu = 100$, $\sigma^2 = 105$, D = {3.38(0), 9.81(0), 10.03(0.001), 37.41(0.522)} respectively.

Table 4.14 shows the results for the Fuzzy c- partition simulation for when n=500, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01$, $\sigma^2 = 0.05$, ,D = {0.62(0), 9.96(0), 10.37(0.007), 34.97(0.476)}. For $\mu = 0.05$, $\sigma^2 = 0.1$, D = {6.50(0), 9.96(0), 11.01(0.019), 36.93(0.513)}. For $\mu = 0.5$, $\sigma^2 = 0.55$ D = {7.39(0), 9.95(0), 10.01(0.000), 34.73(0.471)}. For $\mu = 1$, $\sigma^2 = 1.5$ D = {3.40(0), 6.19(0), 10.15(0.003), 33.10(0.440)}. For $\mu = 2$, $\sigma^2 = 2.5$, D= {8.21(0), 9.97(0), 10.07(0.001), 35.57(0.487)}. For $\mu = 10$, $\sigma^2 = 10.5$, D = {6.99(0), 9.93(0), 10.63(0.012), 37.29(0.520)}. For $\mu = 50$, $\sigma^2 = 55$, D = {6.57(0), 9.88(0), 10.49(0.009), 34.89(0.474)}. For $\mu = 100$, $\sigma^2 = 105$, D = {6.36(0), 9.88(0), 10.06(0.001), 38.46(0.542)} respectively.

Table 4.14 shows the results for the Fuzzy c- partition simulation for when n=1000, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01$, $\sigma^2 = 0.05$, D = {7.48(0), 10.60(0), 34.20(0.011), 37.48(0.461}. For $\mu = 0.05$, $\sigma^2 = 0.1$, D= {11.10(0,021), 32.13(0.422}. For $\mu = 0.5$, $\sigma^2 = 0.55$ D = {8.11(0), 10.82(0.016), 32.14(0.422)}. For $\mu = 1$, $\sigma^2 = 1.5$ D = {8.23(0), 11.73(0.033), 32.64(0.431)}. For $\mu = 2$, $\sigma^2 = 2.5$, D = {10.93(0.018), 32.34(0.426)}. For $\mu = 10, \sigma^2 = 10.5$, D= {10.22(0.004), 31.14(0.403)}. For $\mu = 50$, $\sigma^2 = 55$, D = {7.74(0), 10.62(0.012), 33.09(0.440)}. For $\mu = 100$, $\sigma^2 = 105$, D= {8.76(0), 10.94(0.018), 31.83(0.416)} respectively.

Table 4.15 shows the results for the Fuzzy c- partition simulation for when n=5000, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01$, $\sigma^2 = 0.05$, D = {16.07(0.116), 28.46(0.352)}. For $\mu = 0.05$, $\sigma^2 = 0.1$, D = {16.53(0.124), 27.49(0.333)}. For $\mu = 0.5$, $\sigma^2 = 0.55$ D = {16.07(0.115), 27.77(0.388)}. For $\mu = 1$, $\sigma^2 = 1.5$ D = {15.64(0.107), 27.68(0.337)}. For $\mu = 2$, $\sigma^2 = 2.5$, D = {16.06(0.115), 28.46(0.352)}. For $\mu = 10$, $\sigma^2 = 10.5$, D = {16.53(0.124), 27.49(0.333)}.For $\mu = 50$, $\sigma^2 = 55$, D = {16.07(0.116), 27.77(0.338)}. For $\mu = 100$, $\sigma^2 = 105$, D = {15.64(0.107), 27.68(0.337)} respectively.

Generalised Poisson Model

Table 4.16. Fuzzy set classification of different levels of overdispersion of GP n=20, 30.

n=20

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	0.25 (0)	5.41 (0)	12.44 (0.005)	19.72 (0.185)	44.53 (0.363)	66.40 (0.948)	85.92 (1)	125.00 (1)
0.05	0.1	0.25 (0)	8.82 (0)	11.76 (0.031)	39.02 (0.553)	40.63 (0.580)	65.87 (0.941)	77.47 (1)	98.78 (1)
0.5	0.55	0.75 (0)	9.26 (0)	18.68 (0.165)	39.02 (0.553)	40.41 (0.577)	65.60 (0.937)	70.15 (1)	118.58 (1)
1	1.5	5.56 (0)	-	11.67 (0.032)	38.14 (0.536)	59.90 (0.856)	-	118.58 (1)	146.30 (1)
2	2.5	1.74 (0)	8.82 (0)	11.76 (0.034)	27.76 (0.338)	43.80 (0.626)	50.61 (0.755)	90.00 (1)	-
10	10.5	3.69 (0)	9.30 (0)	11.44 (0.028)	38.27 (0..539)	47.41 (0.677)	69.72 (0.996)	112.39 (1)	117.71 (1)
50	55	1.73 (0)	7.99 (0)	10.47 (0.009)	39.92 (0.570)	52.37 (0.748)	-	86.57 (1)	111.11 (1)
100	105	4.08 (0)	9.31 (0)	13.64 (0.069)	27.98 (0.342)	43.15 (0.616)	65.87 (0.941)	90.00 (1)	-
		n=30							
z		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	0.27 (0)	8.06 (0)	13.74 (0.071)	39.92 (0.570)	46.15 (0.659)	60.79 (0.998)	72.73 (1)	87.32 (1)
0.05	0.1	0.22 (0)	8.16 (0)	11.67 (0.009)	39.92 (0.570)	43.51 (0.622)	63.18 (0.802)	89.43 (1)	-
0.5	0.55	0.27 (0)	9.81 (0)	13.02 (0.058)	37.48 (0.523)	42.20 (0.577)	61.44 (0.996)	81.15 (1)	84.28 (1)
1	1.5	0.85 (0)	9.31 (0)	16.35 (0.121)	38.71 (0.547)	41.95 (0.738)	67.65 (0.966)	80.17 (1)	85.92 (1)
2	2.5	0.00 (0)	9.62 (0)	10.47 (0.009)	29.74 (0.376)	45.89 (0.617)	52.87 (0.755)	-	-
10	10.5	0.84 (0)	9.52 (0)	12.64 (0.050)	37.48 (0.523)	41.84 (0.598)	43.70 (0.624)	81.15 (1)	-
50	55	1.46 (0)	9.62 (0)	10.24 (0.004)	35.71 (0.490)	47.22 (0.622)	53.68 (0.766)	-	-
100	105	0.00 (0)	5.56 (0)	10.62 (0.012)	40.00 (0.571)	40.21 (0.574)	67.36 (0.962)	83.03 (1)	124.16 (1)

10.18(0.003), 37.08(0.516)}. For $\mu = 100$, $\sigma^2 = 105$, D = {3.38(0), 9.81(0), 10.03(0.001), 37.41(0.522)} respectively.

Table 4.16 – 4.19 show the results for the Fuzzy c- partition simulation for GP, Table 4.16 shows the results when n=20, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01, \sigma^2 = 0.05$, D = {0.25(0), 5.41 (0), 12.44 (0.005), 19.72(0.185), 44.53(0.363), 66.40(0.948), 85.92(1), 125.00(1)}. For $\mu = 0.05$, $\sigma^2 = 0.1$, D= {0.25(0), 8.82 (0), 11.76 (0.031), 39.02(0.553), 40.41(0.577), 65.60(0.937), 70.15(1), 118.58 (1)}. For $\mu = 0.5$, $\sigma^2 = 0.55$ D = {0.75 (0), 9.26 (0), 18.68(0.165), 39.02(0.553), 40.41(0.577), 65.60(0.937), 70.15(1), 118.58 (1)}. For $\mu = 1$, $\sigma^2 = 1.5$ D= {5.56 (0), 11.67(0.032), 38.14(0.536), 59.90(0.856), 118.58(1), 146.30 (1)}. For $\mu = 2$, $\sigma^2 = 2.5$, D= {1.74(0), 8.82(0), 11.76(0.034), 27.76 (0.338), 43.80(0.626), 50.61(0.755), 90.00(1)}. For $\mu = 10$, $\sigma^2 = 10.5$, D = {3.69(0), 9.30 (0), 11.44(0.027), 38.27 (0.539), 47.41(0.677), 69.72(0.996), 112.39(1), 117.71(1)}. For $\mu = 50$, $\sigma^2 = 55$, D = {1.73 (0), 7.99 (0), 10.47(0.009), 10.47(0.009), 39.92(0.570), 52.37(0.748), 86.57 (1), 111.11 (1)}. For $\mu = 100$, $\sigma^2 = 105$, D = {4.08(0), 9.31 (0), 13.64(0.069), 27.98 (0.342), 43.15 (0.616), 65.87 (0.941), 90.00 (1)} respectively.

Table 4.16 shows the results for the Fuzzy c- partition simulation for when n=30, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01$, $\sigma^2 = 0.05$, D = {0.27(0), 8.06 (0), 13.74 (0.019), 39.92(0.570), 46.15 (0.659), 60.79(0.998), 72.73(1), 87.32(1)}. For $\mu = 0.05$, $\sigma^2 = 0.1$, D = {0.22(0), 8.16 (0), 11.67 (0.009), 39.92(0.570), 43.51 (0.622), 63.18(0.802), 89.43(1)}. For $\mu = 0.5$, $\sigma^2 = 0.55$ D = {0.27 (0), 9.81 (0), 0.75 (0), 13.02 (0.058), 37.48 (0.523), 42.20 (0.577), 61.44(0.996), 81.15 (1), 84.28(1)}. For $\mu = 1$, $\sigma^2 = 1.5$, D= {0.85 (0), 9.31 (0), 16.35 (0.121), 38.71 (0.547), 41.95 (0.738), 67.65 (0.966), 80.17(1), 85.92(1)}. For $\mu = 2$, $\sigma^2 = 2.5$, D= {0, 9.62 (0), 10.47 (0.009), 29.74 (0.376), 45.89 (0.617), 52.87 (0.755)}. For $\mu = 10$, $\sigma^2 = 10.5$, D = {0.84 (0), 9.52 (0), 12.64 (0.050), 43.70 (0.624), 81.15(1)}. For $\mu = 50$, $\sigma^2 = 55$, D= {1.46 (0), 9.62 (0), 10.24 (0.004), 35.71 (0.490), 47.22(0.622), 53.68 (0.766)}. For $\mu = 100$, $\sigma^2 = 105$, D = {0, 5.56 (0), 10.62 (0.012), 40.21 (0.574), 67.36 (0.962), 83.03(1)} respectively.

Table 4.17. Fuzzy set classification of different levels of overdispersion of GP

n=50, 100

n=50

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	0.00 (0)	9.52 (0)	10.99 (0.019)	39.02 (0.553)	40.57 (0.580)	69.87 (0.998)	71.05 (1)	128.48 (1)
0.05	0.1	0.27 (0)	9.81 (0)	10.47 (0.009)	39.04 (0.553)	40.03 (0.572)	67.36 (0.992)	80.95 (1)	84.33 (1)
0.5	0.55	2.22 (0)	9.62 (0)	10.11 (0.002)	39.36 (0.559)	40.41 (0.577)	65.60 (0.937)	76.85 (1)	160.54 (1)
1	1.5	1.46 (0)	9.81 (0)	10.42 (0.009)	40.00 (0.571)	40.03 (0.572)	66.02 (0.943)	70.15 (1)	111.11 (1)
2	2.5	0.31 (0)	9.81 (0)	10.38 (0.007)	39.92 (0.570)	41.67 (0.595)	69.72 (0.996)	76.36 (1)	171.43 (1)
10	10.5	0.00 (0)	8.71 (0)	10.27 (0.005)	37.84 (0.530)	41.49 (0.593)	65.87 (0.941)	76.85 (1)	149.80 (1)
50	55	0.00 (0)	9.52 (0)	10.27 (0.005)	39.92 (0.570)	43.53 (0.622)	68.35 (0.976)	71.05 (1)	137.50 (1)
100	105	0.00 (0)	9.52 (0)	10.24 (0.004)	40.00 (0.571)	41.14 (0.588)	68.35 (0.976)	80.58 (1)	161.59 (1)
		n=100							
		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	0.00 (0)	9.52 (0)	10.99 (0.019)	39.02 (0.553)	40.57 (0.580)	69.87 (0.998)	71.05 (1)	128.48 (1)
0.05	0.1	0.27 (0)	9.81 (0)	10.47 (0.009)	39.04 (0.553)	40.03 (0.572)	67.36 (0.992)	80.95 (1)	84.33 (1)
0.5	0.55	2.22 (0)	9.62 (0)	10.11 (0.002)	39.36 (0.559)	40.41 (0.577)	65.60 (0.937)	76.85 (1)	160.54 (1)
1	1.5	1.46 (0)	9.81 (0)	10.42 (0.009)	40.00 (0.571)	40.03 (0.572)	66.02 (0.943)	70.15 (1)	111.11 (1)
2	2.5	0.31 (0)	9.81 (0)	10.38 (0.007)	39.92 (0.570)	41.67 (0.595)	69.72 (0.996)	76.36 (1)	171.43 (1)
10	10.5	0.00 (0)	8.71 (0)	10.27 (0.005)	37.84 (0.530)	41.49 (0.593)	65.87 (0.941)	76.85 (1)	149.80 (1)
50	55	0.00 (0)	9.52 (0)	10.27 (0.005)	39.92 (0.570)	43.53 (0.622)	68.35 (0.976)	71.05 (1)	137.50 (1)
100	105	0.00 (0)	9.52 (0)	10.24 (0.004)	40.00 (0.571)	41.14 (0.588)	68.35 (0.976)	80.58 (1)	161.59 (1)

Table 4.17 shows the results for the Fuzzy c- partition simulation for when n=50, the minimum and maximum values for not severe, moderately severe, severe and very severe

For $\mu = 0.05$, $\sigma^2 = 0.1$, D = {0.27(0), 7.64 (0), 10.42 (0.009), 35.71 (0.490), 54.05 (0.772), 67.65 0.966), 83.87(1), 180.08(1)}. For $\mu = 0.5$, $\sigma^2 = 0.55$ D = {0.84 (0), 8.82 (0), 11.44 (0.028), 39.36 (0.559), 42.20 (0.603), 66.02(0.943), 71.53(1), 105.54(1)}. For $\mu = 1$, $\sigma^2 = 1.5$, D = {1.46 (0), 9.81 (0), 11.44 (0.028), 35.75 (0.490), 41.95 (0.599), 69.72(0.996), 80.00 1), 206.87 (1)}. For $\mu = 2$, $\sigma^2 = 2.5$, D= {1.79(0), 9.31 (0), 11.76 (0.034), 39.02 (0.553), 45.89 (0.656), 67.36 (0.962), 72.73(1), 120.22 (1)}. For $\mu = 10$, $\sigma^2 = 10.5$, D = {1.13 (0), 8.31 (0), 10.62 (0.012), 40.00(0.571), 41.84 (0.598), 66.02(0.943), 72.73(1), 140.85 (1)}. For $\mu = 50$, $\sigma^2 = 55$, D= {0.25(0), 8.31 (0), 13.32 (0.063), 40.00(0.571), 47.22(0.675), 69.72(0.996), 78.13(1), 140.85 (1)}. For $\mu = 100$, $\sigma^2 = 105$, D = {0.25(0), 9.62 (0), 10.77 (0.015), 37.70 (0.528), 43.15(0.616), 69.72(0.996), 71.05(1), 98.84 (1)} respectively.

Table 4.17 shows the results for the Fuzzy c- partition simulation for when n=100, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01$, $\sigma^2 = 0.05$, D = {0, 9.52(0), 10.99 (0.019), 39.02(0.553), 40.57 (0.580), 69.87(0.998), 71.05(1), 128.48(1)}. For $\mu = 0.05$, $\sigma^2 = 0.1$, D = {0.27(0), 9.81(0), 10.47(0.009), 39.04(0.553), 40.03 (0.572), 67.36(0.992), 80.95(1), 84.33 (1)}. For $\mu = 0.5$, $\sigma^2 = 0.55$, D= {0, 9.62(0), 10.11(0.002), 39.36 (0.591), 40.41(0.577), 65.60(0.937), 76.85(1), 160.54(1)}. For $\mu = 1$, $\sigma^2 = 1.5$, D = {1.46 (0), 9.81(0), 10.42(0.008), 40.00(0.571), 40.03 (0.572), 66.02(0.996), 70.15(1), 111.11(1)}. For $\mu = 2$, $\sigma^2 = 2.5$, D= {0.31(0), 9.81(0), 10.38(0.007), 39.92(0.570), 41.67 (0.595), 69.72(0.996), 76.36(1), 171.43(1)}. For $\mu = 10$, $\sigma^2 = 10.5$, D= {0, 9.52 (0), 10.27(0.005), 37.84 (0.530),,), 41.49(0.593), 65.87(0.941), 76.85(1), 149.80(1)}. For $\mu = 50$ $\sigma^2 = 55$, D= {0, 9.52 (0), 10.27 (0.005), 39.92(0.570), 43.53(0.580), 68.35(0.976), 71.05(1), 137.50(1)}. For $\mu = 100$ $\sigma^2 = 105$, D= {0, 9.52(0), 10.24(0.004), 40.00(0.571), 41.14(0.588), 68.35(0.976), 80.58(1), 161.59(1)} respectively.

Table 4.18 shows the results for the Fuzzy c- partition simulation for when n=200, the minimum and maximum values for not severe, moderate, severe and very severe

Table 4.18. Fuzzy set classification of different levels of overdispersion of GP

n=200, 300

n=200

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	0.31 (0)	9.62 (0)	10.11 (0.002)	39.21 (0.556)	41.14 (0.588)	67.65 (0.966)	70.15 (1)	182.98 (1)
0.05	0.1	0.00 (0)	9.31 (0)	10.11 (0.002)	39.02 (0.553)	40.11 (0.573)	69.72 (0.996)	70.15 (1)	175.00 (1)
0.5	0.55	0.00 (0)	9.62 (0)	10.11 (0.002)	40.00 (0.571)	41.15 (0.588)	68.35 (0.976)	72.73 (1)	255.71 (1)
1	1.5	0.25 (0)	9.62 (0)	10.11 (0.002)	40.00 (0.571)	40.57 (0.580)	69.72 (0.996)	76.85 (1)	291.18 (1)
2	2.5	0.00 (0)	9.81 (0)	10.77 (0.015	39.84 (0.568)	40.41 (0.577)	69.72 (0.996)	70.15 (1)	287.20 (1)
10	10.5	0.00 (0)	9.81 (0)	10.11 (0.002)	39.92 (0.570)	43.08 (0.615)	66.40 (0.949)	72.73 (1)	140.11 (1)
50	55	0.25 (0)	9.85 (0)	10.47 (0.009)	40.00 (0.571)	40.63 (0.580)	65.87 (0.941)	78.13 (1)	141.82 (1)
100	105	0.00 (0)	9.52 (0)	11.04 (0.020)	40.00 (0.571)	43.18 (0.617)	69.72 (0.996)	72.73 (1)	146.83 (1)

n=300

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	0.00 (0)	9.81 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.72 (0.996)	71.05 (1)	152.53 (1)
0.05	0.1	0.00 (0)	9.62 (0)	10.96 (0.018)	40.00 (0.571)	40.03 (0.572)	69.00 (0.986)	71.05 (1)	241.03 (1)
0.5	0.55	0.00 (0)	9.81 (0)	10.11 (0.002)	39.92 (0.570)	40.57 (0.580)	69.72 (0.996)	72.73 (1)	328.43 (1)
1	1.5	0.00 (0)	9.52 (0)	10.27 (0.005)	40.00 (0.571)	40.16 (0.574)	69.00(0.986)	70.15 (1)	124.16 (1)
2	2.5	0.00 (0)	9.81 (0)	10.11 (0.002)	40.00 (0.571)	40.57 (0.580)	68.35 (0.976)	71.05 (1)	142.55 (1)
10	10.5	0.00 (0)	9.62 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.00 (0.986)	73.94 (1)	375.00 (1)
50	55	0.00 (0)	9.81 (0)	10.11 (0.002)	39.92 (0.570)	40.42 (0.577)	69.72 (0.996)	70.15 (1)	205.93 (1)
100	105	0.22 (0)	9.81 (0)	10.24 (0.004)	39.92 (0.570)	41.11 (0.587)	69.72 (0.996)	71.05 (1)	171.43 (1)

respectively, for $\mu = 0.01$, $\sigma^2 = 0.05, D =$ {0.31(0), 9.62(0), 10.11 (0.002), 39.21(0.556), 41.14(0.588), 67.65(0.966), 70.15(1), 182.98(1)}. For $\mu = 0.05$, $\sigma^2 = 0.1$, D= {0, 9.31(0), 10.11(0.002), 39.02(0.553), 40.11(0.573), 69.72(0.996),70.15(1),

75.00(1)}. For $\mu = 0.5$, $\sigma^2 = 0.55$, D= {0, 9.62(0), 10.11(0.002), 40.00(0.571), 41.15(0.588), 68.35(0.976), 255.71(1)}. For $\mu = 1$, $\sigma^2 = 1.5$, D= {0.25(0), 9.62(0), 10.11(0.002), 40.00(0.571), 40.57(0.580), 69.72(0.996), 70.15(1), 76.85(1), 291.18(1)}. For $\mu = 2$, $\sigma^2 = 2.5$, D= {0, 9.81(0), 10.77(0.015), 39.84(0.568), 40.41(0.577), 69.72(0.996), 70.15(1), 287.20(1)}. 70.15(1), 175.00(1)}. For $\mu = 0.5$, $\sigma^2 = 0.55$, D= {0, 9.62(0), 10.11(0.002), 40.00(0.571), 41.15(0.588), 68.35(0.976), 255.71(1)}. For $\mu = 1$, $\sigma^2 = 1.5$, D= {0.25(0), 9.62(0), 10.11(0.002), 40.00(0.571), 40.57(0.580), 69.72(0.996), 70.15(1), 76.85(1), 291.18(1)}. For $\mu = 2$, $\sigma^2 = 2.5$, D= {0, 9.81(0), 10.77(0.015), 39.84(0.568), 40.41(0.577), 69.72(0.996), 70.15(1), 287.20(1)}. For $\mu = 10$, $\sigma^2 = 10.5$, D= {0, 9.81(0), 10.11(0.002), 39.92(0.570), 43.08(0.615), 66.40(0.949), 72.73(1), 140.11(1)}. For $\mu = 50$, $\sigma^2 = 55$, D= {0.25(0), 9.85(0), 10.47(0.009), 8.31 (0), 13.32 (0.063), 40.00(0.571), 40.63(580), 65.87(0.941), 78.13(1), 141.82(1)}. For $\mu = 100$, $\sigma^2 = 105$, D= {0, 9.52 (0), 11.04(0.020), 40.00(0.571), 43.18(0.617), 69.72(0.996), 146.83(1)} respectively.

Table 4.18 shows the results for the Fuzzy c- partition simulation for when n=300, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01$, $\sigma^2 = 0.05$, D = {0, 9.81(0), 10.11 (0.002), 40.00 (0.571), 40.03 (0.572), 69.72 (0.996), 71.05(1), 152.53(1)}. For $\mu = 0.05, \sigma^2 = 0.1$, D = {0, 9.62(0), 10.96(0.018), 40.00 (0.571), 40.03 (0.572), 69.00 (0.986), 71.05(1), 241.03(1)}. For $\mu = 0.5$, $\sigma^2 = 0.55$, D = {0, 9.81(0), 10.11(0.002), 39.92 (0.570), 40.57 (0.580), 69.72 (0.996), 72.73(1), 328.43(1)}. For $\mu = 1$, $\sigma^2 = 1.5$, D = {0), 9.52(0), 10.27(0.005), 40.00(0.571), 40.16 (0.574), 69.00(0.986), 70.15(1), 124.16(1)}. For $\mu = 2$, $\sigma^2 = 2.5$, D = {0, 9.81(0), 10.11(0.002), 40.00(0.571), 40.57 (0.580), 68.35 (0.976), 71.05(1), 142.55(1)}. For $\mu = 10$, $\sigma^2 = 10.5$, D = {0, 9.62(0), 9.81(0), 10.11(0.002), 40.00(0.571), 40.03 (0.572), 69.00(0.986), 40.42 (0.577), 73.94(1), 375.00(1)}. For $\mu = 50$, $\sigma^2 = 55$, D = {0, 9.81(0), 10.11(0.002), 39.92 (0.570), 69.72 (0.996), 70.15(1), 205.93 1)}. For $\mu = 100$, $\sigma^2 = 105$, D = {0.22(0), 9.81(0), 10.24(0.004), 39.92 (0.570), 41.11(0.587), 69.72(0.996), 71.05(1), 171.43(1)} respectively.

Table 4.19 shows the results for the Fuzzy c- partition simulation for when n=500, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01$, $\sigma^2 = 0.05$, D = {0, 9.81(0), 10.42 (0.009), 40.00 (0.571), 40.15(0.574), 68.17 (0.974), 70.15(1), 180.08(1)}. For $\mu = 0.05$, $\sigma^2 = 0.1$, D = {0, 9.81(0), 10.16(0.003), 40.00 (0.571), 40.41 (0.577), 69.72 (0.996), 70.15(1), 255.78(1)}. For $\mu = 0.5$, $\sigma^2 = 0.55$ D = {0, 9.85(0), 10.16(0.003), 39.92 (0.570), 40.57 (0.580), 68.17 (0.974), 70.11(1), 205.93(1)}. For $\mu = 1$, $\sigma^2 = 1.5$, D = {0), 9.81(0), 10.11(0.002), 40.00(0.571), 40.36 (0.577), 69.72(0.996), 70.15(1), 286.37 (1)}. For $\mu = 2$, $\sigma^2 = 2.5$, D = {0, 9.81(0), 10.14(0.003), 39.36 (0.559), 40.21 (0.574), 69.00(0.986), 71.05(1), 205.36(1)}. For $\mu = 10$, $\sigma^2 = 10.5$, D= {0, 9.85(0), 10.27(0.005), 40.00(0.571), 40.63 (0.580), 69.72(0.996), 70.15(1), 205.36(1)}. For $\mu = 50$. $\sigma^2 = 55$, D= {0, 9.62(0), 10.11(0.002), 40.00(0.571), 69.72 (0.996), 76.36(1), 171.43(1)}. For $\mu = 100$, $\sigma^2 = 105$, D = {0, 9.52(0), 10.11(0.002), 40.00(0.571), 40.03(0.572), 69.72(0.996), 70.15(1), 280.00 (1)} respectively.

Table 4.19 shows the results for the Fuzzy c- partition simulation for when n=1000, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively,For $\mu = 0.01$, $\sigma^2 = 0.05$, ,D = {0, 9.85(0), 10.11 (0.002), 40.00 (0.571), 40.03(0.572), 69.72 (0.996) 70.15(1), 362.82(1)}. For $\mu = 0.05$, $\sigma^2 = 0.1$, D = {0, 9.81(0), 10.38(0.007), 39.97 (0.571), 40.57 (0.580) 69.72 (0.996) 70.15(1), 231.75 (1)}. For $\mu = 0.5$, $\sigma^2 = 0.55$, D = {0, 9.81(0), 10.11(0.002), 40.00 (0.571), 40.03(0.572), 69.87 (0.998, 70.51(1), 216.67(1)}. For $\mu = 1$, $\sigma^2 = 1.5$ D = {0), 9.81(0), 10.11(0.002), 40.00(0.571), 40.03 (0.572), 69.72(0.996), 71.15(1), 189.13(1)}. For $\mu = 2$, $\sigma^2 = 2.5$, D = {0, 9.85(0), 10.14(0.003), 40.00(0.571), 40.03 (0.572), 69.72(0.996), 71.51(1), 216.67(1)}. For $\mu = 10$, $\sigma^2 = 10.5$, D = {0, 9.81(0), 10.11(0.002), 40.00(0.571), 40.03 (0.572), 40.00(0.571), 69.72(0.996), 70.15(1), 320.49(1)}. For $\mu = 50$, $\sigma^2 = 55$, D = {0, 9.62(0), 10.27(0.005), 40.00(0.571), 40.03(0.572) 69.72 (0.996), 70.15(1), 245.45(1)}. For $\mu = 100$, $\sigma^2 = 105$, D= {0, 9.8(0), 10.11(0.002), 40.00(0.571), 40.03(0.572), 69.72(0.996), 71.05(1), 258.20 (1)} respectively. For $\mu = 0.01$, $\sigma^2 = 0.05$, ,D = {0, 9.85(0), 10.11 (0.002), 40.00 (0.571), 40.03(0.572), 69.72 (0.996) 70.15(1), 362.82(1)}.

Table 4.19. Fuzzy set classification of different levels of overdispersion of GP

n=500, 1000, 5000

n=500

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	0.00 (0)	9.81 (0)	10.42 (0.009)	40.00 (0.571)	40.15 (0.574)	68.17 (0.974)	70.15 (1)	180.08 (1)
0.05	0.1	0.00 (0)	9.81 (0)	10.16 (0.003)	40.00 (0.571)	40.41 (0.577)	69.72 (0.996)	70.15 (1)	255.78 (1)
0.5	0.55	0.00 (0)	9.85 (0)	10.16 (0.003)	39.92 (0.570)	40.57 (0.580)	68.17 (0.974)	70.11 (1)	205.93 (1)
1	1.5	0.00 (0)	9.81 (0)	10.11 (0.002)	40.00 (0.571)	40.36 (0.577)	69.72 (0.996)	70.15 (1)	286.37 (1)
2	2.5	0.00 (0)	9.81 (0)	10.14 (0.003)	39.36 (0.559)	40.21 (0.574)	69.00 (0.986)	71.05 (1)	205.36 (1)
10	10.5	0.00 (0)	9.85 (0)	10.27 (0.005)	40.00 (0.571)	40.63 (0.580)	69.72 (0.996)	70.15 (1)	205.36 (1)
50	55	0.00 (0)	9.62 (0)	10.11 (0.002)	40.00 (0.571)	40.16 (0.574)	69.72 (0.996)	76.36 (1)	171.43 (1)
100	105	0.00 (0)	9.52 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.72 (0.996)	70.15 (1)	280.00 (1)

n=1000

		Not Severe		Moderate-ly Severe		Severe		Very Severe	
μ	σ^2	Min		Min	Max	Min	Max	Min	Max
0.01	0.05	0.00 (0)	9.85 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.72 (0.996)	70.15 (1)	362.82 (1)
0.05	0.1	0.00 (0)	9.81 (0)	10.38 (0.007)	39.97 (0.571)	40.57 (0.580)	69.72 (0.996)	70.15 (1)	231.75 (1)
0.5	0.55	0.00 (0)	9.81 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.87 (0.998)	70.51 (1)	216.67 (1)
1	1.5	0.00 (0)	9.81 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.72 (0.996)	71.05 (1)	189.13 (1)
2	2.5	0.00 (0)	9.85 (0)	10.14 (0.003)	40.00 (0.571)	40.03 (0.572)	69.72 (0.996)	70.51 (1)	216.67 (1)

10	10.5	0.00 (0)	9.81 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.72 (0.996)	70.15 (1)	320.49 (1)
50	55	0.00 (0)	9.62 (0)	10.27 (0.005)	40.00 (0.571)	40.03 (0.572)	69.72 (0.996)	70.15 (1)	245.45(1)
100	105	0.00 (0)	9.81 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.72(0.996)	71.05 (1)	258.20 (1)
n=5000									
		Not Severe		Moderately Severe		Severe		Very Severe	
μ	σ^2	Min	Max	Min	Max	Min	Max	Min	Max
0.01	0.05	0.00 (0)	9.85 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.87 (0.998)	70.15 (1)	298.39 (1)
0.05	0.1	0.00 (0)	9.85 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.87 (0.998)	70.15 (1)	328.43 (1)
0.5	0.55	0.00 (0)	9.85 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.87 (0.998)	70.15 (1)	342.47 (1)
1	1.5	0.00 (0)	9.85 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.87 (0.998)	70.15 (1)	271.74 (1)
2	2.5	0.00 (0)	9.85 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.87 (0.998)	71.05 (1)	280.00 (1)
10	10.5	0.00 (0)	9.85 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.87 (0.998)	70.11 (1)	338.46 (1)
50	55	0.00 (0)	9.85 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.87 (0.998)	70.15 (1)	323.74 (1)
100	105	0.00 (0)	9.85 (0)	10.11 (0.002)	40.00 (0.571)	40.03 (0.572)	69.87 (0.998)	70.15 (1)	347.06 (1)

D = {0, 9.81(0), 10.38(0.007), 39.97 (0.571), 40.57 (0.580) 69.72 (0.996) 70.15(1), 231.75 (1)}. For $\mu = 0.5$, $\sigma^2 = 0.55$, D = {0, 9.81(0), 10.11(0.002), 40.00 (0.571), 40.03(0.572), 69.87 (0.998, 70.51(1), 216.67(1)}. For $\mu = 1$, $\sigma^2 = 1.5$ D = {0), 9.81(0), 10.11(0.002), 40.00(0.571), 40.03 (0.572), 69.72(0.996), 71.15(1), 189.13(1)}.

For $\mu = 2$, $\sigma^2 = 2.5$, D = {0, 9.85(0), 10.14(0.003), 40.00(0.571), 40.03 (0.572), 69.72(0.996), 71.51(1), 216.67(1)}. For $\mu = 10$, $\sigma^2 = 10.5$, D = {0, 9.81(0), 10.11(0.002), 40.00(0.571), 40.03 (0.572), 40.00(0.571), 69.72(0.996), 70.15(1), 320.49(1)}. For $\mu = 50$, $\sigma^2 = 55$, D = {0, 9.62(0), 10.27(0.005), 40.00(0.571), 40.03(0.572) 69.72 (0.996), 70.15(1), 245.45(1)}. For $\mu = 100$, $\sigma^2 = 105$, D= {0, 9.8(0), 10.11(0.002), 40.00(0.571), 40.03(0.572), 69.72(0.996), 71.05(1), 258.20 (1)} respectively.

Table 4.19 shows the results for the Fuzzy c- partition simulation for when n=5000, the minimum and maximum values for not severe, moderately severe, severe and very severe respectively, for $\mu = 0.01, \sigma^2 = 0.05,$,D = {0, 9.85(0), 10.11 (0.002), 40.00 (0.571), 40.03(0.572), 69.87(0.998), 70.15(1), 298.39(1)}. For $\mu = 0.05, \sigma^2 = 0.1$, D = {0, 9.85(0), 10.11 (0.002), 40.00 (0.571), 40.03(0.572), 69.87(0.998), 70.15(1), 328.43(1)}. For $\mu = 0.5$, $\sigma^2 = 0.55$, D = {0, 9.85(0), 10.11 (0.002), 40.00 (0.571), 40.03(0.572), 69.87(0.998), 70.15(1), 342.47(1)}. For $\mu = 1$, $\sigma^2 = 1.5$, D = {0, 9.85(0), 10.11 (0.002), 40.00 (0.571), 40.03(0.572), 69.87(0.998), 70.15(1), 271.74(1)}. For $\mu = 2$, $\sigma^2 = 2.5$, D = {0, 9.85(0), 10.11 (0.002), 40.00 (0.571), 40.03(0.572), 69.87(0.998), 70.05(1), 280.00(1)}. For $\mu = 10$, $\sigma^2 = 10.5$, D = {0, 9.85(0), 10.11 (0.002), 40.00 (0.571), 40.03(0.572), 69.87(0.998), 70.11(1), 338.46 (1)}. For $\mu = 50$, $\sigma^2 = 55$, D = {0, 9.85(0), 10.11 (0.002), 40.00 (0.571), 40.03(0.572), 69.87(0.998), 70.15(1), 323.74(1)}. For $\mu = 100$, $\sigma^2 = 105$, D = {0, 9.85(0), 10.11 (0.002), 40.00 (0.571), 40.03(0.572), 69.87(0.998), 70.15(1), 347.06 (1)} respectively.

4.2. Averaging Method for Determination of Threshold

The averaging method was used for the determination of the threshold in this research. The average of each of the dispersion percentages was computed to determine the threshold. Table 20-23 are the threshold values for each of the sample sizes that is, the minimum values and the maximum values. At this point, the minimum and maximum threshold At this point, the minimum and maximum threshold values are presented values are presented in the Table with different dispersion percentages and sample sizes. The minimum values refer to the point the model should begin to be considered for modification however, the minimum threshold can still be tolerated while the maximum values indicate the point in which the model should be considered for modification.

Table 4.20 shows that for sample size n=20 for the Poisson model the minimum threshold value is 16.78% and the maximum threshold value is 34.07% which means for the sample size n=20 Poisson model should be modified when the threshold is 34.07%. The minimum threshold value for the Negative Binomial is 69.35% and the maximum threshold value is 78.43%,. For Com-Poisson model the minimum threshold value is 19.33% and the maximum threshold value is 54.69%; and for Generalised Poisson model the minimum threshold value is 23.75% and the maximum threshold value is 37.09%. At the maximum threshold values, the following models should be considered for modification.

Table 4.20 shows that for sample size n=30, for the Poisson model the minimum threshold value is 20.20% and the maximum threshold value is 30.75% which means for the sample size n=30 Poisson model should be modified when the threshold is 30.75%. The minimum threshold value for the Negative Binomial is 69.85% and the maximum threshold value is 77.08%. For Com-Poisson model the minimum threshold value is 17.89% and the maximum threshold value is 26.95%; and for Generalised Poisson model the minimum threshold value is 16.75% and the maximum threshold value is 36.48%.

Table 4.20 shows that for sample size n=50, for the Poisson model the minimum threshold value is 24.85% and the maximum threshold value is 32.19 % which means for the sample size n=50, Poisson model should be modified when the threshold is 32.19%. The minimum threshold value for the Negative Binomial is 72.96% and the maximum threshold value is 76.55%. For Com-Poisson model, the minimum threshold value is 20.57% and the maximum threshold value is 25.12%, and for Generalised Poisson model the minimum threshold value is 25.51% and the maximum threshold value is 36.55%. At the maximum threshold values, the following models should be considered for modification Table 4.20 shows that for sample size n=50, for the is

32.19 % which means for the sample size n=50, Poisson model should be modified when the threshold is 32.19%. The minimum threshold value for the Negative Binomial is 72.96% and the maximum threshold value is 76.55%.

For Com-Poisson model, the minimum threshold value is 20.57% and the maximum threshold value is 25.12%, and for Generalised Poisson model the minimum threshold value is 25.51% and the maximum threshold value is 36.55%. At the maximum threshold values, the following models should be considered for modification

Table 4.21 shows that for sample size n=100, for the Poisson model the minimum threshold value is 23.80% and the maximum threshold value is 30.76% which means for the sample size n=100 Poisson model should be modified when the threshold is 30.76%. The minimum threshold value for the Negative Binomial is 73.17% and the maximum threshold value is 75.46%. For Com-Poisson model, the minimum threshold value is 19.98 % and the maximum threshold value is 22.62%, and for Generalised Poisson model the minimum threshold value is 24.15% and the maximum threshold value is 31.59%. At the maximum threshold values, the following models should be considered for modification.

Table 4.21 shows that for sample size n=200, for the Poisson model the minimum threshold value is 26.65% and the maximum threshold value is 29.09% which means, for the sample size n=200, the Poisson model should be modified when the threshold is 29.09%. The minimum threshold value for the Negative Binomial is 74.14% and the maximum threshold value is 76.01%. For Com-Poisson model, the minimum threshold value is 21.14% and the maximum threshold value is 22.12%, and for Generalised Poisson model the minimum threshold value is 24.83% and the maximum threshold value is 32.45%. At the maximum threshold values, the following models should be considered for modification.

Table 4.21 shows that for sample size n=300, for the Poisson model the minimum threshold value is 25.70% and the maximum threshold value is 29.95% which means for the sample size n=300 Poisson model should be modified when the threshold is 29. 95%. The minimum threshold value for the Negative Binomial is 73.97% and the maximum threshold value is 75.98%. Com-Poisson model, the minimum threshold va-

Table 4.20. Averaging Method for Determination of Threshold n=20, 30, 50

n=20

μ	σ^2	Poisson	Negative	Com-Poisson	GP
0.01	0.05	24.17	69.35	34.76	32.58
0.05	0.1	26.61	73.57	26.53	27.14
0.5	0.55	16.78	76.31	24.94	34.36
1	1.5	23.41	75.17	25.10	37.09
2	2.5	30.44	77.51	54.69	24.50
10	10.5	34.07	76.79	19.33	35.12
50	55	19.32	75.35	25.49	28.70
100	105	20.86	78.43	30.46	23.75

n=30

μ	σ^2	Poisson	Negative	Com-Poisson	GP
0.01	0.05	25.08	77.08	26.52	36.48
0.05	0.1	24.33	73.39	21.92	28.61
0.5	0.55	21.36	69.85	20.54	28.94
1	1.5	29.19	76.76	26.95	24.59
2	2.5	30.50	72.98	17.89	16.75
10	10.5	20.20	75.40	25.62	21.39
50	55	29.34	73.70	24.02	18.88
100	105	30.75	73.34	25.52	30.42

n=50

μ	σ^2	Poisson	Negative	Com-Poisson	GP
0.01	0.05	26.37	76.55	25.05	30.44
0.05	0.1	26.69	72.96	22.56	33.73
0.5	0.55	26.03	74.23	22.30	32.50
1	1.5	25.93	74.71	25.12	36.55
2	2.5	32.19	74.24	20.57	29.69
10	10.5	24.85	76.49	24.86	32.74
50	55	29.70	73.21	21.56	27.86
100	105	26.88	74.92	25.00	25.51

Table 4.21. Averaging Method for Determination of Threshold, n=100, 200, 300

n=100

μ	σ^2	Poisson	Negative	Com-Poisson	GP
0.01	0.05	27.98	73.89	21.42	27.29
0.05	0.1	25.64	74.85	20.85	24.15
0.5	0.55	30.76	73.17	20.95	30.48
1	1.5	23.80	73.25	22.62	30.79
2	2.5	26.05	75.00	20.47	31.70
10	10.5	25.59	74.96	21.55	24.15
50	55	29.46	75.46	19.98	31.59
100	105	24.99	75.36	21.84	26.93
n=200					
μ	σ^2	Poisson	Negative	Com-Poisson	GP
0.01	0.05	28.45	75.65	22.10	28.28
0.05	0.1	27.38	75.79	21.40	27.19
0.5	0.55	27.44	74.14	21.15	29.10
1	1.5	28.45	74.94	21.14	32.45
2	2.5	28.88	74.61	21.48	28.59
10	10.5	29.09	74.63	21.38	24.83
50	55	27.26	76.01	22.12	25.68
100	105	26.65	74.57	22.11	29.28
n=300					
μ	σ^2	Poisson	Negative	Com-Poisson	GP
0.01	0.05	26.99	75.00	21.52	27.12
0.05	0.1	27.66	74.90	21.56	28.56
0.5	0.55	29.95	75.84	21.8	28.50
1	1.5	28.18	74.80	21.79	24.53
2	2.5	26.84	75.98	22.17	30.10
10	10.5	28.78	73.97	21.68	27.85
50	55	25.70	75.09	21.85	28.22
100	105	26.52	75.17	21.66	27.82

lue is 21.80% and the maximum threshold value is 22.17%, and for Generalised Poisson model the minimum threshold value is 24.53% and the maximum threshold value is 30.10%. At the maximum threshold values, the following models should be considered for modification.

Table 4.22 shows that for sample size n=500. For the Poisson model the minimum threshold value is 26.43% and the maximum threshold value is 28.90% which means, for the sample size n=500, Poisson model should be modified when the threshold is 28.90%. The minimum threshold value for the Negative Binomial is 74.26% and the maximum threshold value is 75.49%. For Com-Poisson model, the minimum threshold value is 21.49% and the maximum threshold value is 22.28%, and for Generalised Poisson model, the minimum threshold value is 27.09% and the maximum threshold value is 30.32%. At the maximum threshold values, the following models should be considered for modification.

Table 4.22 shows that for sample size n=1000, for the Poisson model, the minimum threshold value is 27.06% and the maximum threshold value is 29.51% which means for the sample size n=1000, Poisson model should be modified when the threshold is 29.51%. The minimum threshold value for the Negative Binomial is 27.60% and the maximum threshold value is 75.28%. For Com-Poisson model, the minimum threshold value is 31.14% and the maximum threshold value is 34.20%, and for Generalised Poisson model the minimum threshold value is 27.26% and the maximum threshold value is 30.34%. At the maximum threshold values, the following models should be considered for modification.

Table 4.22 shows that for sample size n=5000, For the Poisson model the minimum threshold value is 27.60% and the maximum threshold value is 28.24% which means for the sample size n=5000, Poisson model should be modified when the threshold is 28.24%. The minimum threshold value for the Negative Binomial is 29.05% and the maximum threshold value is 74.86%. For Com-Poisson model the minimum threshold value is 22.13% and the maximum threshold value is 22.20%, and for Generalised Poisson model, the minimum threshold value is 28.21% and the maximum threshold value is 29.20%. At the maximum threshold values, the following models should be considered for modification.

Table 4.23 shows for sample size n = 20, 30… 5000, the minimum and the maximum threshold value for the P0, NB, CP, and GP model respectively. When n= 20, the minimum threshold value for PO, NB, CP, and GP is 16.78%, 69.35%, 19.33% and 23.75% while the maximum threshold value is 34.07%, 78.43%, 54.69% and 37.09% respectively. When n= 30, the minimum threshold value for PO, NB, CP, and GP is 20.2%, 69.85%, 17.89%, and 16.75% while the maximum threshold value is 30.75%, 77.08%, 26.52% and 36.4%8 respectively.

When n= 50, the minimum threshold value for PO, NB, CP, and GP is 24.85%, 72.96%, 20.57% and 25.51% while the maximum threshold value is 26.88%, 76.55%, 25.12% and 33.73%. When n= 100, the minimum threshold value for PO, NB, CP, and GP is 23.80%, 73.17%, 19.98% and 24.15% while the maximum threshold value is 30.76%, 75.46%, 21.48% and 31.70% respectively. When n= 200, the minimum threshold value for PO, NB, CP, and GP is 26.65%, 74.14%, 21.14% and 24.83% while the maximum threshold value is 29.09%, 76.01%, 22.12% and 34.45% respectively. When n= 300, the minimum threshold value for PO, NB, CP, and GP is 25.70%, 73.97%, 21.52% and 24.53% while the maximum threshold value is 29.95%, 75.98%, 22.17% and 30.10% respectively.

When n= 500, the minimum threshold value for PO, NB, CP, and GP is 26.43%, 74.26%, 21.49% and 27.09% while the maximum threshold value is 28.25%, 75.18%, 22.12% and 30.32% respectively. When n= 1000, the minimum threshold value for PO, NB, CP, and GP is 27.06%, 27.56%, 31.14% and 27.26% while the maximum threshold value is 29.51%, 75.28%, 34.20% and 30.34% respectively. When n= 5000, the minimum threshold value for PO, NB, CP, and GP is 27.60%, 29.05%, 74.86%, 22.13%, and 28.21% while the maximum threshold value is 28.24%, 74.86%, 22.20% and 29.20% respectively. The largest sample size n=5000 were used to determine the threshold for the Modification of the four-count models. Table 4.24 shows for sample size n = 5000, the minimum and the maximum threshold value for the P0, NB, CP, and GP model is 27.60%, 28.24%, 29.05%, 74.86%, 22.13%, 22.20%, 28.21% and 29.20% respectively.

4.3. Application to Accident Data

In this research, the Poisson regression model, Negative Binomial regression model, Com-Poisson, and Generalised Poisson model were used to model the incidence of

Road traffic crashed in Nigeria. The study is based on the secondary data from the Bulletin of the Federal Road Safety Corps of Nigeria between 2014- 2018. A cross-sectional data of incidence of Road crash in Nigeria in all the thirty-six states and Federal Capital Territory (FCT) of Nigeria were collected from the Bulletin of Federal Road Safety Corps of Nigeria between 2014 and 2018. The models are commonly used ones for count data and numbers of road crashes are good examples of count data. Six variables were used for this analysis; they are the number of crashes (Y) which is also the count SPV=Speed violation (SPV), Using of Phone while driving (UPD), Overloading (OVL), Dangerous driving (DGD), and Sleeping while driving (SOS) are the independent variables. Usually, the test of multicollinearity is carried out to determine if there is a violation of the basic assumptions of the classical linear regression model. Table 4.25 is the result of the multicollinearity test using the Tolerance and Variance Inflation Factor to verify the assumption of the model.

4.4 MULTICOLLINEARITY

In regression analysis, a multicollinearity test is usually conducted to ensure that there is no violation of one of the basic assumptions of the linear regression model to have a reliable result of the study. In the research, the test is conducted for the independent variable to ensure there is no multicollinearity among the explanatory variables. The Variance inflation factor and tolerance are conducted for the five explanatory variables The rule of thumb is that when the variance inflation factor (VIF) that is VIF >10 then, collinearity or multicollinearity exists among the variables. As it is shown from Table 4.25 none of the explanatory variable variance inflation factor is greater than 10, this means that multicollinearity does not exist among the variables and can therefore be used for modeling the count data and can also be used for further study of accident data.

4.6 Parameter Estimation and Statistical Inference

The parameters of the model were estimated to study the impact of the covariates on the study. Table 4.27 – 4.30 present the results for the estimation and the figure 4.1-4.10 for additional information. Figure 4.1 to Figure 4.6 show that the distribution of the accident data.

Table 4.27 shows the result of the analysis of the Road accident data for Poisson mod-

Table 4.22. Averaging Method for Determination of Threshold n=500, 1000, 5000

n=500

μ	σ^2	Poisson	Negative	Com-Poisson	GP
0.01	0.05	26.68	75.18	22.12	27.37
0.05	0.1	28.26	74.26	21.84	29.65
0.5	0.55	28.90	74.83	21.87	27.09
1	1.5	27.42	74.44	22.11	30.32
2	2.5	28.25	74.80	21.95	28.35
10	10.5	26.78	75.49	22.28	28.08
50	55	26.43	74.74	21.70	27.39
100	105	27.77	74.98	21.49	29.31
n=1000					
μ	σ^2	Poisson	Negative	Com-Poisson	GP
0.01	0.05	28.23	74.89	34.20	28.40
0.05	0.1	27.80	75.28	32.13	27.26
0.5	0.55	27.06	74.36	32.14	30.34
1	1.5	27.17	75.24	32.64	28.00
2	2.5	27.76	29.22	32.34	28.32
10	10.5	28.77	27.60	31.14	29.25
50	55	29.51	32.45	33.09	29.77
100	105	28.55	42.19	31.83	29.68
n=5000					
μ	σ^2	Poisson		Com-Poisson	GP
0.01	0.05	27.67	74.86	22.13	28.33
0.05	0.1	27.60	74.53	22.20	28.92
0.5	0.55	27.63	37.17	22.16	28.21
1	1.5	27.91	38.16	22.15	29.20
2	2.5	28.24	29.05	22.17	28.28
10	10.5	28.12	35.35	22.16	28.92
50	55	28.21	41.37	22.19	28.47
100	105	27.96	44.12	22.17	28.42

Table 4.23. Summary of the minimum and maximum threshold values for the four count models.

Model/Sample sizes	Poisson		NB		CP		GP	
	Min	max	Min	Max	Min	Max	Min	Max
20	16.78	34.07	69.35	78.43	19.33	54.69	23.75	37.09
30	20.2	30.75	69.85	77.08	17.89	26.52	16.75	36.48
50	24.85	26.88	72.96	76.55	20.57	25.12	25.51	33.73
100	23.8	30.76	73.17	75.46	19.98	21.84	24.15	31.70
200	26.65	29.09	74.14	76.01	21.14	22.12	24.83	32.45
300	25.70	29.95	73.97	75.98	21.52	22.17	24.53	30.10
500	26.43	28.25	74.26	75.18	21.49	22.12	27.09	30.32
1000	27.06	29.51	27.56	75.28	31.14	34.20	27.26	30.34
5000	**27.60**	**28.24**	**29.05**	**74.86**	**22.13**	**22.20**	**28.21**	**29.20**

Table 4.24. The Threshold values for the Models

Model	Minimum	Maximum
Poisson	27.60	28.24
Negative Binomial	29.05	74.86
Com-Poisson	22.13	22.20
Generalised Poisson	28.21	29.20

Table 4.25. Collinearity statistics

Model	Tolerance	VIF
SPV	0.697	1.434
UPD	0.801	1.249
OVL	0.752	1.327
DGD	0.813	1.229
SOS	0.972	1.029

4.5 Exploratory Data Analysis

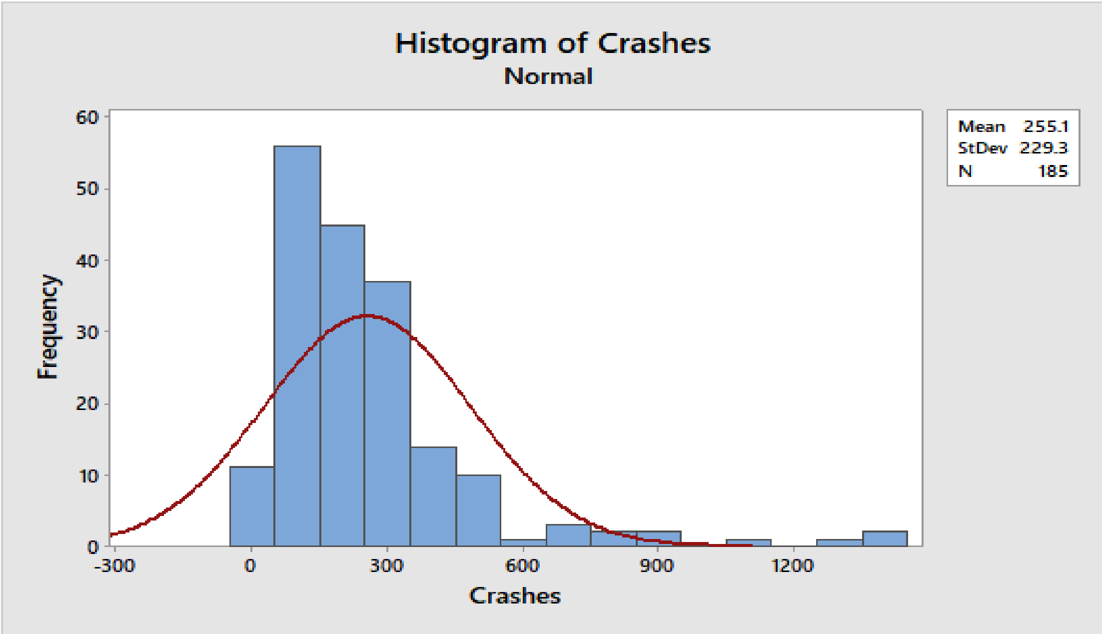

Figure 4.1. Plot for number of crashes

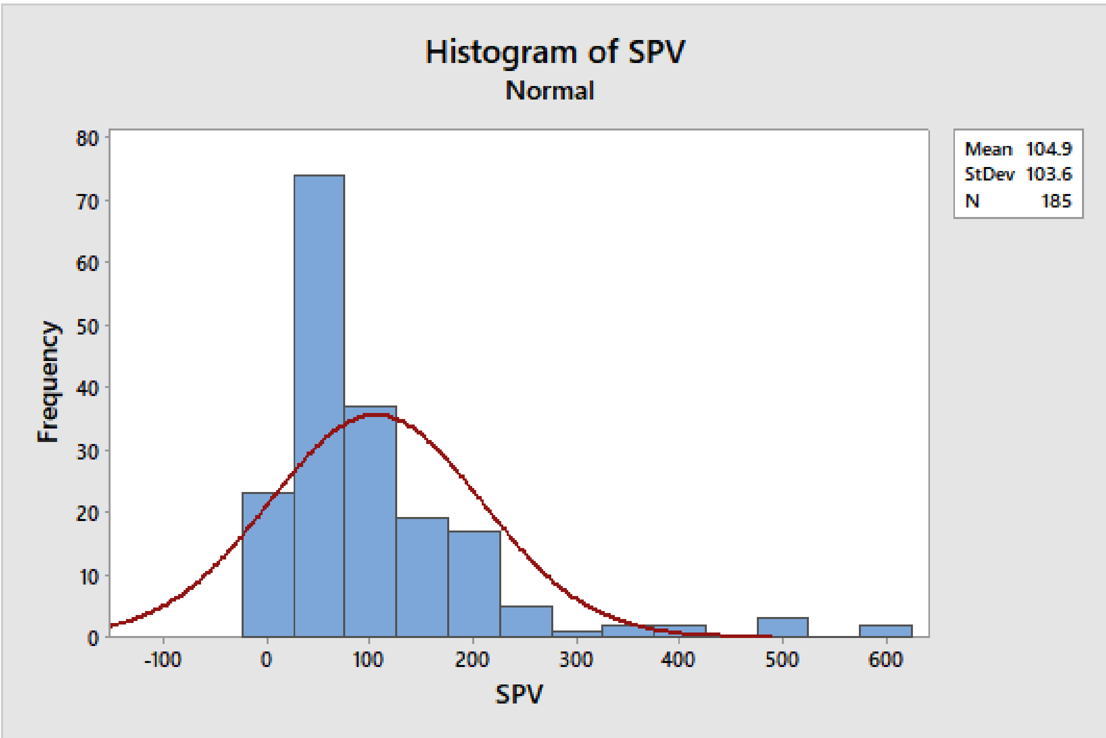

Figure 4.2. Plot for number of SPV

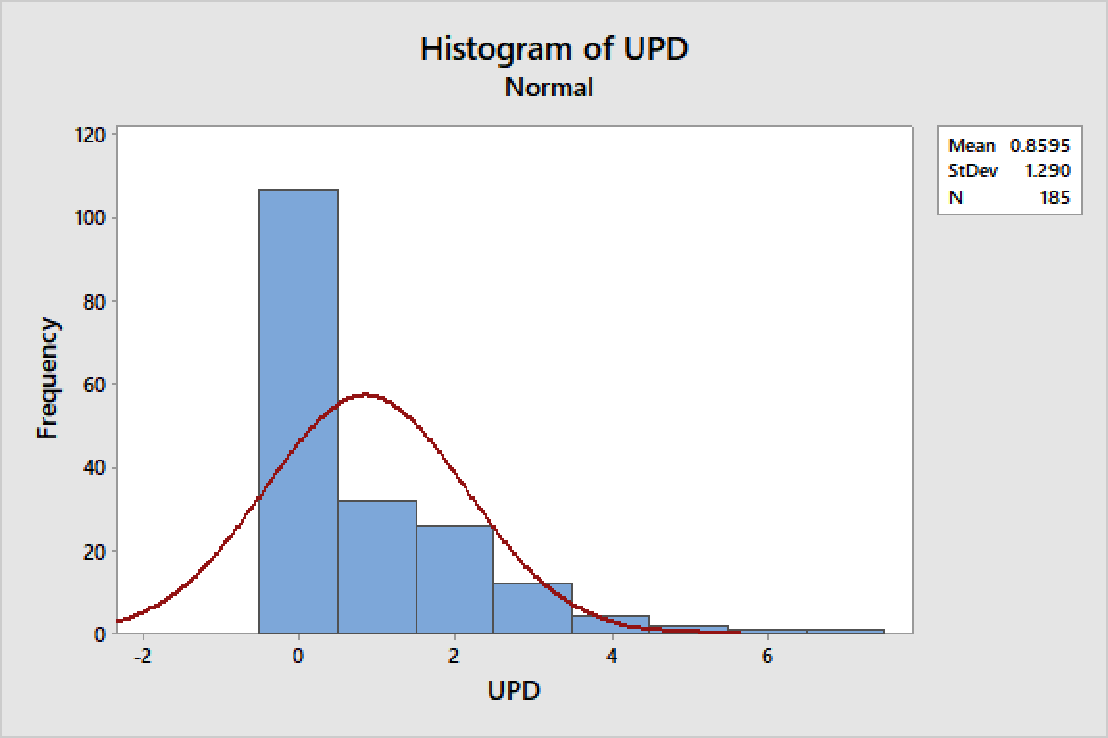

Figure 4.3. Plot for number of UPD

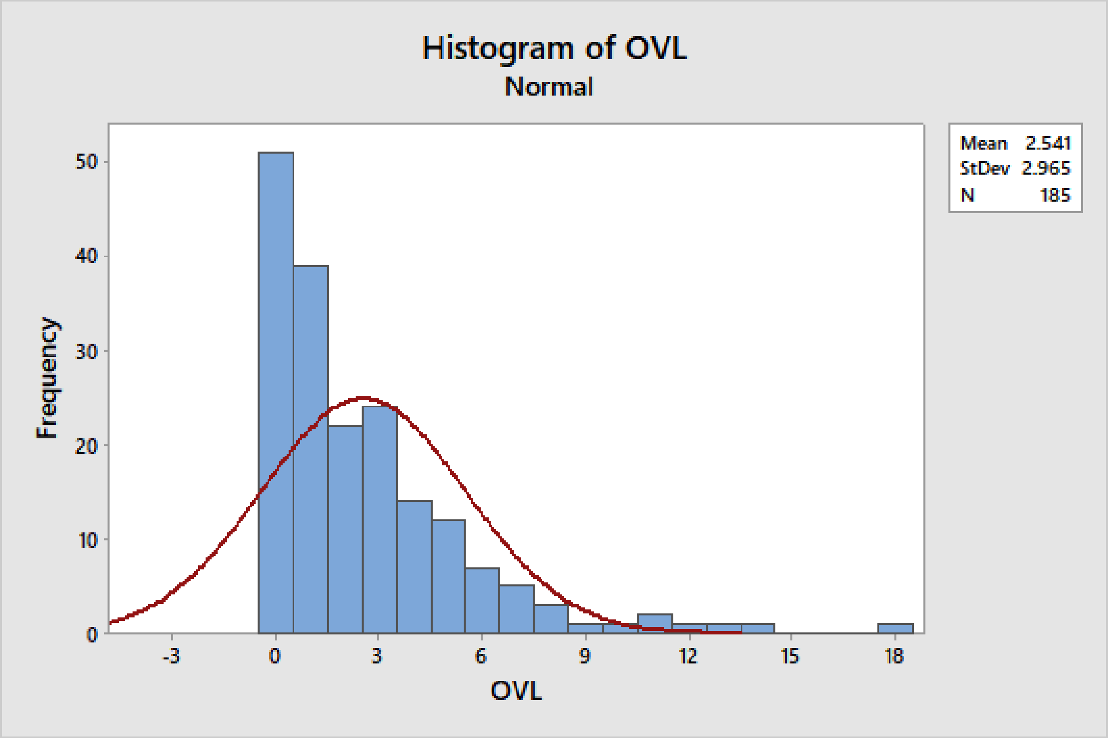

Figure 4.4. Plot for number of OVL

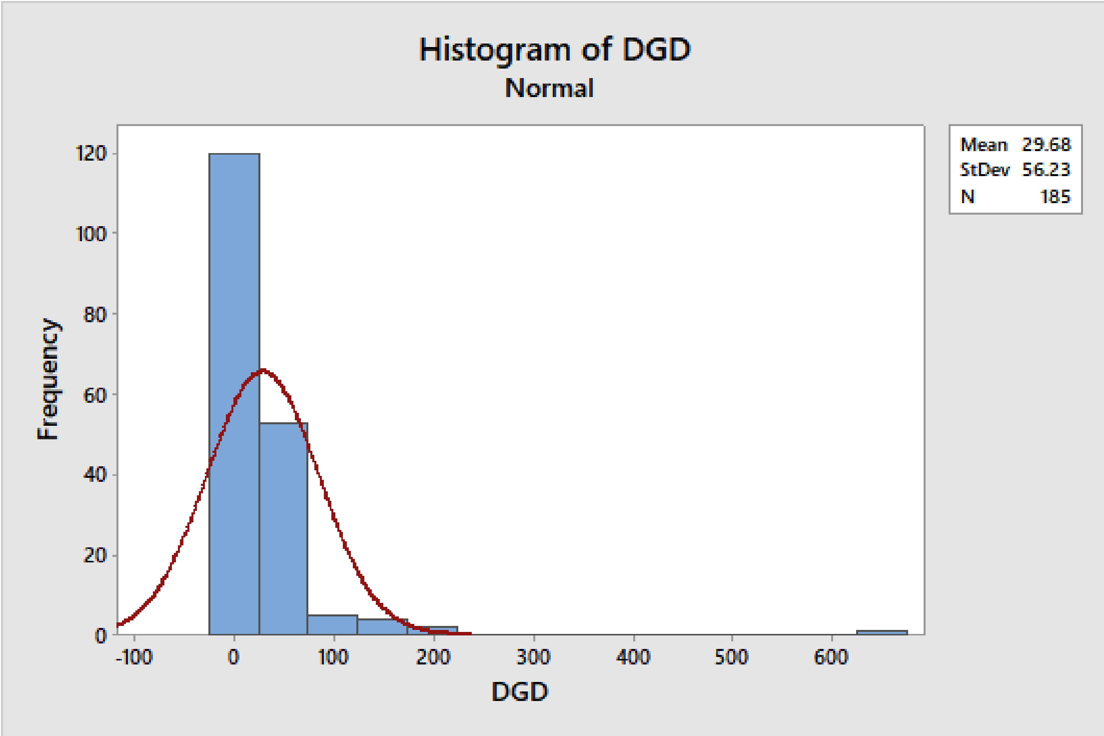

Figure 4.5. Plot for number of DGD.

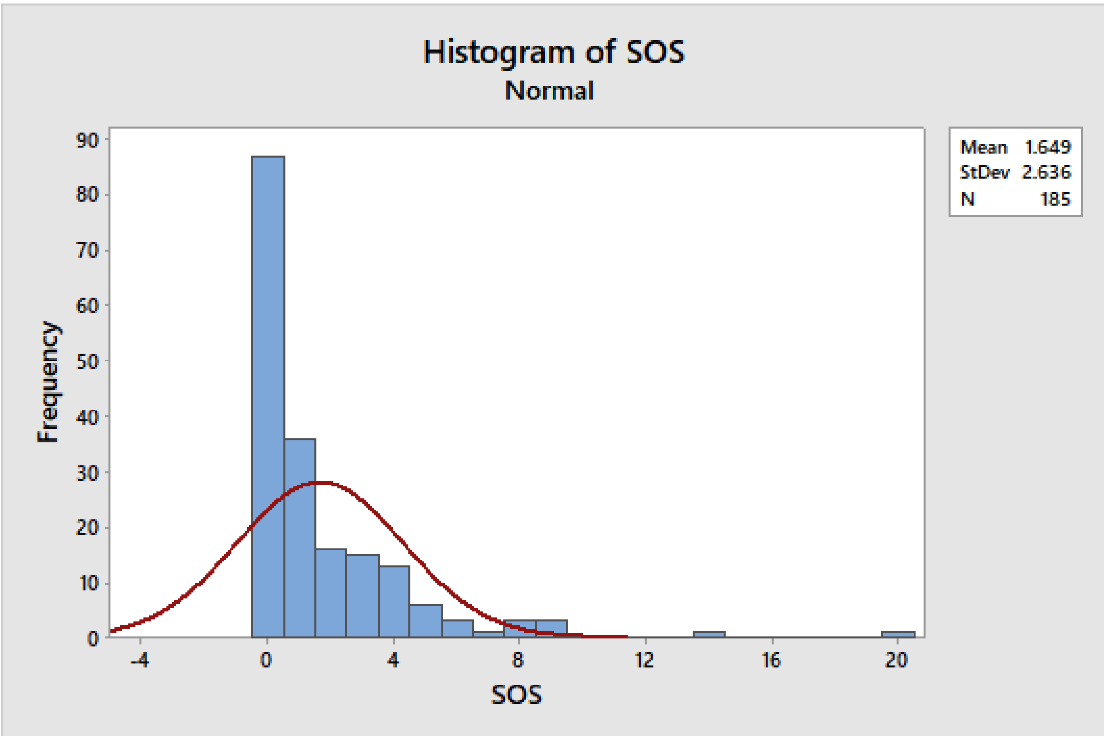

Figure 4.6. Plot for number of SOS

el. The estimate of the intercept and of the coefficient of SPV, UPD, OVL, DGD, and SOS is 4.786, 3.568x10^{-3}, 5.172x10^{-2}, 3.271x10^{-2}, 1.146x10^{-3}, 4.786, and 3.042x10^{-2} respectively. The estimate of standard error for the intercept and for the coefficient of SPV, UPD, OVL, DGD, and SOS is 8.152x10^{-3}, 3.334x10^{-5}, 3.274x10^{-3}, 1.502x10^{-3}, 6.588x10^{-5}, and 1.715x10^{-3} respectively. The score for the intercept and the explanatory variables is 587.0700, 106.6900, 17.4500, 21.7900, 17.3900 and 17.7400. The P-value for each of the coefficients is 2x10^{-16}. The P-value is compared to the level of significance which is 5% for this research. It is revealed from the analysis that the intercept and the explanatory variable used to model the data are significant.

Table 4.28 shows the estimation of Negative Binomial model for the intercept and of the coefficient of SPV, UPD, OVL, DGD, and SOS is 4.6292, 0.0043, 0.0140, 0.0408, 0.0015 and 0.0270 respectively. The estimate of standard error for the intercept and for the coefficient of SPV, UPD, OVL, DGD, and SOS is 0.0558, 0.0004, 0.0303, 0.0136, 0.0007 and 0.0136 respectively. The Normal score for the intercept and the explanatory variables is 82.9294, 11.9830, 0.4622, 3.0069, 2.1892 and 1.9846. The P-value for each of the coefficients is 2x10^{-16}. The P-value is compared to the level of significance which is 5% for this research. It is revealed from the analysis that the intercept and the explanatory variables used to model the data are significant.

Table 4.29. shows the estimation of Com-Poisson model for the intercept and of the coefficients of SPV, UPD, OVL, DGD, and SOS is 4.7722, 0.0036, 0.0515, 0.0329, 0.0011 and 0.038 respectively. The estimate of standard error for the intercept and for the coefficient of SPV, UPD, OVL, DGD, and SOS is 0.4784, 0.0040, 0.0054, 0.0031, 0.0001 and 0.0044 respectively. The Normal score for the intercept and the explanatory variables is 9.9747, 9.2976, 9.5533, 10.7090, 9.9900 and 8.7119 respectively. The P-value for intercept and of the coefficient of SPV, UPD, OVL, DGD, and SOS is 1.966x10^{-23}, 1.436x10-^{20}, 1.257x10^{-21}, 9.234x$10^{-27,}$ 1.687x10^{-23}, and 2.989x10^{-18} respectively. The P-value is compared to the level of significance which is 5% for this research. It is revealed from the analysis that the intercept and the explanatory variable used to model the data are significant.

Table 4.30. shows the estimation of Generalised Poisson model of the intercept and of the coefficient of SPV, UPD, OVL, DGD, and SOS is 14.7516, 0.0036 , 0.0630,

0.0367, 0.0010, and 0.0372 respectively. The estimate of standard error for the intercept and for the coefficient of SPV, UPD, OVL, DGD, and SOS is 0.0563, 0.0002, 0.0217, 0.0098, 0.0005 and 0.0108 respectively. The Normal score for the intercept and the explanatory variables is 84.3590, 15.5730, 2.9010, 3.7560, 2.1960 and 3.4320 respectively. The P-value for intercept and of the coefficient of SPV, UPD, OVL, DGD, and SOS is 1.966×10^{-23}, 1.436×10^{-20}, 2×10^{-16}, 2×10^{-16}, 0.0037, 0.0002, 0.0281 and 0.0006 respectively. The P-value is compared to the level of significance which is 5% for this research. It is revealed from the analysis that the intercept and the explanatory variable used to model the data are significant.

Figure 4.7-4.11. is the Quantile-Quantile (QQ) plot for PO, NB, CP and GP to compare with Normal Distribution. It reveals from the plots that the points are not on the straight line. If all the points fall on the straight line, it is shows that the distribution is normally distributed.

4.7 Model Selection

Model selection is necessary whenever different models are used for data set to propose the best model that properly fit the data sets. The Akaike Information Criterion (AIC) is used to select the best model to fit the data in this research. The AIC values for the models were presented in Table 4.31. The model with the least value has a better fit which is the most appropriate to fit the data. Table 4.31 shows the result of the AIC for the four models used to fit the road crash accident data to know the most appropriate model to fit the accident data. The AIC for PO, NB, CP and GP is 8836.70, 2211.03, 8657 and 2205.42 respectively. GP has the least value.

Multi-collinearity test was carried out using the Variance Inflation Factor (VIF) to know if there is a violation of the basic assumption of the classical linear regression model before fitting the count with the different models. From the results, it is revealed that none of the values is greater than 10 because when the VIF is greater than 10 there is evidence of collinearity. The count data was first fitted with the Poisson model which showed evidence of overdispersion, The p values at ($p < 0.005$) shows that all the variables used for the models are significant at 5% level of significance which means all the variables contributed to the causes of the road accident and these should

Table 4.26. Fuzzy set methods for classifying the Accident data different percentages of overdispersion.

Variable	Mean	Variance	dispersion	Group	Membership
Crashes	255.10	52,579.23	99.51	Very Severe	1

Table 4.27. Parameter Estimation of Poisson Regression Model

Parameter	Estimate	Standard error	Z-value	P-value
Intercept	4.786	$8.152x10^{-3}$	587.0700	$2x10^{-16}$
SPV	$3.568x10^{-3}$	$3.334x10^{-5}$	106.6900	$2x10^{-16}$
UPD	$5.172x10^{-2}$	$3.274x10^{-3}$	17.4500	$2x10^{-16}$
OVL	$3.271x10^{-2}$	$1.502x10^{-3}$	21.7900	$2x10^{-16}$
DGD	$1.146x10^{-3}$	$6.588x10^{-5}$	17.3900	$2x10^{-16}$
SOS	$3.042x10^{-2}$	$1.715x10^{-3}$	17.7400	$2x10^{-16}$

Table 4.28. Parameter Estimation of Negative Binomial Regression model

Parameter	Estimate	Standard error	Z-value	P-value
Intercept	4.6292	0.0558	82.9294	$2x10^{-16}$
SPV	0.0043	0.0004	11.9830	$2x10^{-16}$
UPD	0.0140	0.0303	0.4622	$2x10^{-16}$
OVL	0.0408	0.0136	3.0069	$2x10^{-16}$
DGD	0.0015	0.0007	2.1892	$2x10^{-16}$
SOS	0.0270	0.0136	1.9846	$2x10^{-16}$

Table 4.29. Parameter Estimation of Com-Poisson Regression Model

Parameter	Estimate	Standard error	Z-value	P-value
Intercept	4.7722	0.4784	9.9747	$1.966x10^{-23}$
SPV	0.0036	0.0040	9.2976	$1.436x10^{-20}$
UPD	0.0515	0.0054	9.5533	$1.257x10^{-21}$
OVL	0.0329	0.0031	10.7090	$9.234x10^{-27}$
DGD	0.0011	0.0001	9.9900	$1.687x10^{-23}$
SOS	0.0381	0.0044	8.7119	$2.989x10^{-18}$

Table 4.30. Parameter Estimation of Generalised Poisson Regression Model

Parameter	Estimate	Standard error	Z-value	P-value
Intercept	14.7516	0.0563	84.3590	$2x10^{-16}$
SPV	0.0036	0.0002	15.5730	$2x10^{-16}$
UPD	0.0630	0.0217	2.9010	0.0037
OVL	0.0367	0.0098	3.7560	0.0002
DGD	0.0010	0.0005	2.1960	0.0281
SOS	0.0372	0.0108	3.4320	0.0006

Table 4.31. Akaike Information Criteria (AIC)

Model	AIC
Poisson	8836.70
Negative Binomial	2211.03
Com- Poisson	8657.64
General Poisson	2205.42

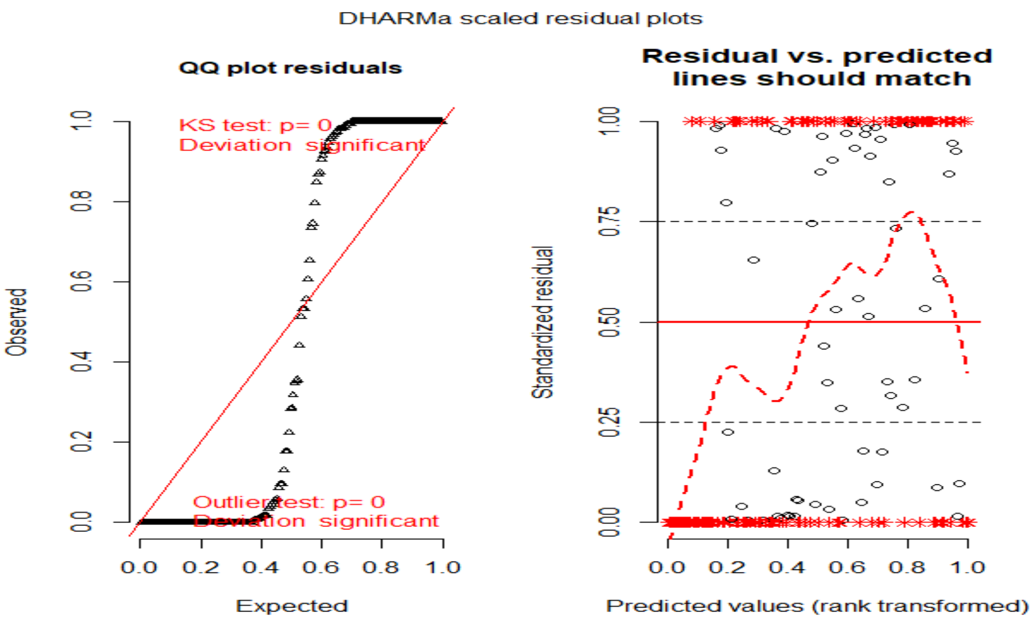

Figure 4.7. QQ- Plot for Poisson Model

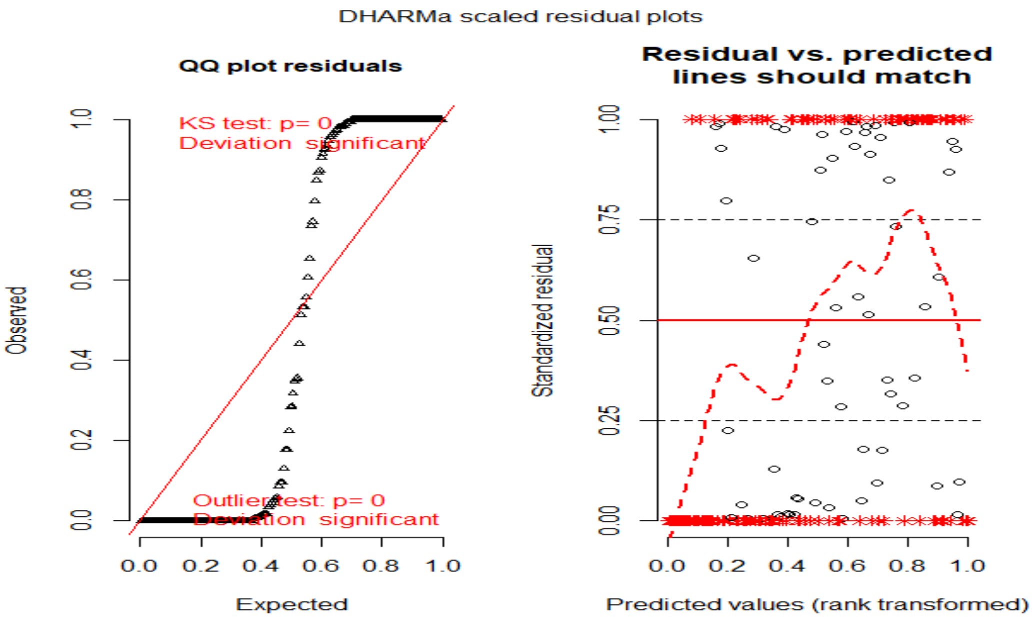

Figure 4.8. QQ- Plot for Negative Binomial Model.

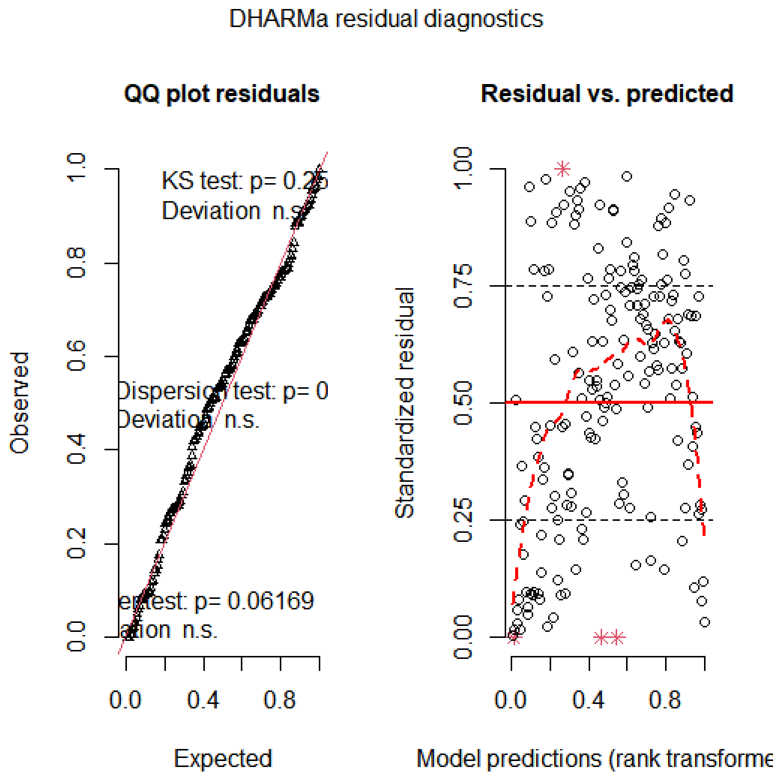

Figure 4.9. QQ-Plot for Com-Poisson Model.

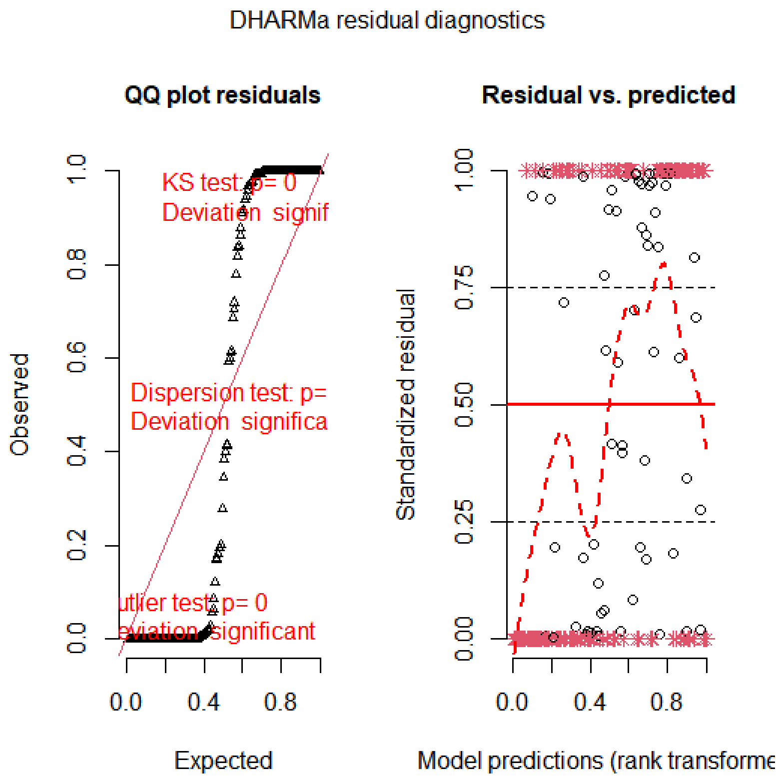

Figure 4.10. QQ-Plot for Generalised Poisson Model.

be included in further study of the road accident crashes.

The count data was first modeled with the Poisson model which showed the evidence of overdispersion. The residual of deviance to its degree of freedom was examined, the rule of thumb is that whenever the ratio of residual deviance to its degree of freedom is equal to one it means that overdispersion does not exist in the data, the data is properly fitted by the Poisson Model but if the ratio of residual of the freedom to its degree of freedom is less than one, there is under dispersion but it is problem of overdispersion when the ratio of residual of the freedom to its degree of freedom greater than one. The residual deviance is 7519.70, the degree of freedom is 179, the ratio of the residual deviance to the degree of freedom is 42.01 which is greater than one. This is a scenario of overdispersion. At this point, alternative models are considered.

4.8 Discussion of Results

In the literature, model have been developed to solve the problem of overdispersion this has been discussed in literature review. Models have been modified and extensively discussed but the threshold before modifying models when the problem of overdispersion inherent in the data was not considered. In this research, different models for count data have been considered with the threshold when needed to modify the models when the need arises. A fuzzy set has been used to classify the different levels of overdispersion, which are: Not severe, Moderate, severe, and very severe which is based on membership function. This has been applied to the life study of road crashes accident between the years 2014 and 2018. Different dispersion percentages result from the simulation study in order to propose the threshold to modify model in the problem of overdispersion is not properly taking care of , the output of the result of simulation will be make available on request. In this study, Fuzzy-c was used to categorise the degree of severity namely Not severe, Moderate, severe, and very severe with their corresponding membership function.. A dispersion percentage of (0 - 10%) is classified as not severe with the membership function of 0. dispersion percentage of (11 - 30%) as moderately severe, (30 - 69%) as severe while 70% and above as Very severe.

The thresholds proposed fills the gap in research in sense that count data models were modified without considering when overdispersion is of serious cause of concern. Famoye (2004) modified Poisson model when the sample mean and sample variance were 0.76 and 1.32 respectively, in line with this study in Table 4.1-4.9 the first column is mean –variance when there is problem of overdispersion as pointed out by Consul et al (1992, Greene (1994) , Ismail et al (2013). Trivedi and Cameroon (1986) modified Poisson model when the sample mean was 0.302 and sample variance was 0.637,but Trivedi and Cameroon (1986) did not consider the threshold before modifying the model. Shmuelli (2005) used an alternative model for the data when the sample mean and sample variance were 3.56 and 11.31 respectively, Lambert(1992) encountered too many zeros (75% of zeros) when modeling number of defect in manufacturing, another model was employed to fit the data. In relationship to this research work.

Table 4.1-4.5 presented the summary of different dispersion percentages of overdispersion of the simulation study of Poission model. In Table 4.1, the minimum value for not severe is 0 and the maximum value is 9.81 with a corresponding membership function of 0, the moderate has the minimum value of 10.11 with membership function of 0.02 and maximum value of 12.95 corresponding membership function of 0.056, severe with the minimum dispersion percentage of 40.36 with the membership function of 0.577 and maximum value of 69.72 with membership function of 0.996 and very severe has a minimum dispersion percentage of 83.03 with membership function of 1 and the maximum value of 116 with membership function of 1. The class of not severe takes a minimum of 0 with a corresponding membership function 0 and maximum takes a value of 9.52 with membership function of 0 in Table 4.1 while moderate has a minimum dispersion percentage 10.11 with 0.002 as the membership function and maximum value of 39.92 with membership function of 0.570, severe has a value of 40 .03 as the minimum and 68.35 with membership function of 0.976 as the maximum value while very severe has a minimum of 71.05 and maximum value with membership function of 1.

Table 4.2 has minimum and maximum value of 0 and 9.81 with membership function of 0 for not severe, 10.11 and 40.00 for minimum and maximum value with membership function of 0.002 and 0.571 respectively for moderate, severe has a value of 40.57 with membership function of 0.580 and maximum value 61 70 with membership function of 0.881 while very severe has a minimum of 71 05 and 169.13 with membership function of 1. In Table 4.2, not severe has a minimum of 0 and maximum value of 9.81 with membership function of 0 , moderate has a minimum of

10.24 and maximum value of 11.44 with membership function of 0.004 and 0.028 respectively, severe has a minimum 40.03 has a maximum value of 66.79 with membership function of 0.572 and 0.954 respectively, while very severe has a minimum of 97.17 and 171.43 with membership function of 1.

The class of not severe takes a minimum of 0 with a corresponding membership function 0 and maximum takes a value of 9.25 with membership function of 0 in Table 4.3 while moderate has a minimum dispersion percentage 10.11 with 0.002 and maximum value 40.00 with the membership function of 0.571 , severe has a value of 40 .16 as the minimum and 69.72 with membership function of 0.996 as the maximum value while very severe has a minimum of 70.15 and maximum value of 210 with membership function of 1. Table 4.3 has minimum and maximum value of 0 and 9.81 with membership function of 0 for not severe, 10.11 and 40.00 for minimum and maximum value with membership function of 0.002 and 0.571 respectively for moderate, severe has a value of 40.03 with membership function of 0.572 and maximum value 69.72 with membership function of 0.996 while very severe has a minimum of 70.15 and 259 with membership function of 1.

Table 4.4 has minimum and maximum value of 0 and 9.81 with membership function of 0 for not severe, 10.11 and 40.00 for minimum and maximum value with membership function of 0.002 and 0.571 respectively for moderate, severe has a value of 40.03 with membership function of 0.572 and maximum value 69.72 with membership function of 0.996 while very severe has a minimum of 71 05 and 298 with membership function of 1. In Table 4.4, not severe has a minimum of 0 and maximum value of 9.61 with membership function of 0 , moderate has a minimum of 10.01 and maximum value of 40.00 with membership function of 0.002 and 0.0571 respectively, severe has a minimum 40.03 has a maximum value of 66.72 with membership function of 0.572 and 0.996 respectively, while very severe has a minimum of 70.15 and 287 with membership function of 1.

The class of not severe takes a minimum of 0 with a corresponding membership function 0 and maximum takes a value of 9.85 with membership function of 0 in Table 4.5, while moderate has a minimum dispersion percentage 10.01 with 0.002 and maximum value 40.00 with the membership function of 0.571 , severe has a value of 40 .03 as the minimum and 69.87 with membership function of 0.998 as the maximum value while very severe has a minimum of 70.15 and maximum value of 411.54 with membership function of 1.

Table 4.6 – 4.10 presented the summary of different dispersion percentages of Over-dispersion of the simulation study of Negative Binomial Model. Table 4.6 Class of Severe has a minimum value of 41.49 with membership function of 0.593 and maximum value 69.67 with membership function of 0.956 while very severe has a minimum of 70.28 and 91.04 with membership function of 1. Table 4.6 , 32.40 and 34.89 for minimum and maximum value with membership function of 0.408 and 0.474 respectively for moderate, severe has a value of 66.27 with membership function of 0.947 and maximum value 69.98 with membership function of 0.977 while very severe has a minimum of 70.28 and 91.04 with membership function of 1.

In Table 4.7, class of severe has a minimum 41.52 has a maximum value of 69.90 with membership function of 0.593 and 0.999 respectively, while very severe has a minimum of 70.16 and 92 75 with membership function of 1. In Table 4.7, moderate has a minimum dispersion percentage 11.99 with 0.038 and maximum value 39.68 with the membership function of 0.565 , severe has a value of 41.90 as the minimum and 69.91 with membership function of 0.99 as the maximum value while very severe has a minimum of 70.01 and maximum value of 94.28 with membership function of 1.

Table 4.8, has the minimum and maximum value of 16.53 and 35.90 with membership function of 0.124 and 0.493 respectively for moderate, severe has a value of 40.20 with membership function of 0.574 and maximum value 69.69 with membership function of 0.999 while very severe has a minimum of 70.02 and 93.57 with membership function of 1. Table 4.8 has 18.36 and 36.31 for minimum and maximum value with membership function of 0.159 and 0.699 respectively for moderate, severe has a value of 40.11 with membership function of 0.573 and maximum value 69.99 with membership function of 0.999 while very severe has a minimum of 70.00 and 95.27 with membership function of 1.

Table 4.9, 15.99 and 39.78 for minimum and maximum value with membership function of 0.114 and 0.567 respectively for moderate, for the class of severe has a value of 40.35 with membership function of 0.666 and maximum value 69.97 with membership function of 0.999 while very severe has a minimum of 70.00 and 95.96

with membership function of 1. Table 4.9, has minimum and maximum value of 0 and 9.99 with membership function of 0 for not severe, 10.00 and 40.00 for minimum and maximum value with membership function of 0.002 and 0.571 respectively for moderate, severe has a value of 40.03 with membership function of 0.572 and maximum value 69.99 with membership function of 0.999 while very severe has a minimum of 70. 05 and 249.88 with membership function of 1. In Table 4.10, not severe has a minimum of 0 and maximum value of 10.00 with membership function of 0 , moderate has a minimum of 10.01 and maximum value of 40.00 with membership function of 0.002 and 0.0571 respectively, severe has a minimum 40.00 has a maximum value of 69.99 with membership function of 0.571 and 0.999 respectively, while very severe has a minimum of 70.00 and 171.43 with membership function of 1.

Table 4.11 - 4.15 presented the summary of different dispersion percentages of overdispersion of the simulation study of Com-Poission model. In Table 4.11 the minimum value for not severe is 0.09 and the maximum value is 9.64 with a corresponding membership function of 0, the moderate has the minimum value of 10.08 with membership function of 0.02 and maximum value of 39.94 corresponding membership function of 0.570, severe with the minimum dispersion percentage of 40.16 with the membership function of 0.574 and maximum value of 69.30 with membership function of 0.990 and very severe has a minimum dispersion percentage of 70.15 and the maximum value of 86.22 with membership function of 1.

The class of not severe takes a minimum of 0.40 with a corresponding membership function 0 and maximum takes a value of 9.94 with membership function of 0 in Table 4.11 while moderate has a minimum dispersion percentage 10.21 with 0.004 as the membership function and maximum value of 39.42 with membership function of 0.566, severe has a value of 40 .44 as the minimum and maximum value of 67.20 with membership function of 0.578 and 0.960 respectively; while very severe has a minimum of 74.05 and the maximum value is greater than 91.9 with membership function of 1.

Table 4.12, has minimum and maximum value of 0.01 and 9.99 with membership function of 0 for not severe, 10.01 and 39.88 for minimum and maximum value with membership function of 0.001 and 0.569 respectively for moderate, severe has a value of 40.20 with membership function of 0.574 and maximum value 53.50 with membership function of 0.764 while very severe has a minimum of 72.94 and the maximum value is greater than 72.94 with membership function of 1. In Table 4.12, not severe has a minimum of 0.14 and maximum value of 9.85 with membership function of 0 , moderate has a minimum of 10.00 and maximum value of 39.74 with membership function of 0 and 0.567 respectively, severe has a minimum value of 40.19 and a maximum value of 47.74 with membership function of 0.574 and 0.682 respectively.

The class of not severe takes a minimum of 0.03 with a corresponding membership function 0 and maximum takes a value of 9.98 with membership function of 0 in Table 4.13, while moderate has a minimum dispersion percentage 10.05 with 0.001 and maximum value 39.91 with the membership function of 0.570 , severe has a value of 40 .10 as the minimum with membership function of 0.587. Table 4.13, has minimum and maximum value of 0.43 and 9.91 with membership function of 0 for not severe, 10.05 and 38.95 for minimum and maximum value with membership function of 0.001 and 0.511 respectively for moderate. Table 4.14, has minimum and maximum value of 0.62 and 9.93 with membership function of 0 for not severe, 10.01 and 38.46 for minimum and maximum value with membership function of 0.001 and 0.522 respectively for moderate. Table 4.14, has minimum and maximum value of 0.43 and 10.60 with membership function of 0 for not severe, 10.22 and 37.48 for minimum and maximum value with membership function of 0.004 and 0.461 respectively for moderate. Table 4.15 has minimum and maximum value of 15.64 and 28.46 for moderate with membership function 0.107 and 0.352 respectively.

Table 4.16- 4.19 presented the summary of different dispersion percentages of overdispersion of the simulation study of Generalised Poisson Model. In Table 4.16 the minimum value for not severe is 0.25 and the maximum value is 9.82 with a corresponding membership function of 0, the moderate has the minimum value of 10.47 with membership function of 0.09 and maximum value of 39.92 with a corresponding membership function of 0.570, severe with the minimum dispersion percentage of 40.41 with the membership function of 0.577 and maximum value of 69.72 with membership function of 0.996 and very severe has a minimum dispersion percentage of 85.92 with membership function of 1 and the maximum value of 146.30 with membership function of 1. In Table 4.16 , not severe has a minimum of 0

and maximum value of 9.81 with membership function of 0 , moderate has a minimum of 10.24 and maximum value of 40.00 with membership function of 0.004 and 0.0571 respectively, severe has a minimum value of 40.21 and a maximum value of 66.72 with membership function of 0.574 and 0.966 respectively.

Table 4.17, 0.25 and 9.81 for minimum and maximum value with membership function of 0 for not severe, 10.42 and 40.00 for minimum and maximum value with membership function of 0.009 and 0.571 respectively for moderate, for the class of severe , a minimum value of 41.95 with membership function of 0.599 and maximum value 69.72 with membership function of 0.996 while very severe has a minimum of 83.87 and 206.87 with membership function of 1. Table 4.17 has minimum and maximum value of 0 and 9.81 with membership function of 0 for not severe, 10.11 and 40.00 for minimum and maximum value with membership function of 0.002 and 0.571 respectively for moderate, for class of severe, a value of 40.03 with membership function of 0.572 and maximum value 69.87 with membership function of 0.998 while very severe has a minimum of 71. 05 and 171.43 with membership function of 1.

In Table 4.18, not severe has a minimum of 0 and maximum value of 9.85 with membership function of 0 , moderate has a minimum of 10.11 and maximum value of 40.00 with membership function of 0.002 and 0.0571 respectively, severe has a minimum 40.11 and maximum has a value of 69.72 with membership function of 0.573 and 0.996 respectively, while very severe has a minimum value of 70.15 and 291 with membership function of 1. Table 4.18 has minimum and maximum value of 0 and 9.85 with membership function of 0 for not severe, 10.11 and 40.00 for minimum and maximum value with membership function of 0.002 and 0.571 respectively for moderate, severe has a value of 40.03 with membership function of 0.572 and maximum value 69.72 with membership function of 0.996 while very severe has a minimum of 70.15 and 328 with membership function of 1.

Table 4.19 has minimum and maximum value of 0 and 9.81 with membership function of 0 for not severe, 10.11 and 40.00 for minimum and maximum value with membership function of 0.002 and 0.571 respectively for moderate, severe has a value of 40.15 with membership function of 0.574 and maximum value 69.72 with membership function of 0.996 while very severe has a minimum of 70.15 and 286.37 with membership function of 1.

In Table 4.19 , not severe has a minimum of 0 and maximum value of 9.61 with membership function of 0 , moderate has a minimum of 10.01 and maximum value of 40.00 with membership function of 0.002 and 0.0571 respectively, severe has a minimum 40.03 has a maximum value of 66.72 with membership function of 0.572 and 0.996 respectively, while very severe has a minimum of 70.15 and 287 with membership function of 1. The class of not severe takes a minimum of 0 with a corresponding membership function 0 and maximum takes a value of 9.85 with membership function of 0 in Table 4.19 while moderate has a minimum dispersion percentage 10.11 with 0.002 and maximum value 40.00 with the membership function of 0.571 , severe has a value of 40 .03 as the minimum and 69.72 with membership function of 0.996 as the maximum value while very severe has a minimum of 70.05 and maximum value of 362 with membership function of 1.

The study has filled the gap in research, in line with Famoye (2004), Trivedi and Cameron (1986), Shmeulli (2004), Consul et'al (1992) polnted out that overdispersion occurs when mean and variance are not equal and different models have been developed. And the first column of Table 4.1-4.19 in this chapter is a case of overdispersion. Means are 0.01, 005,0.5, 1, 2, 10, 50 and 100 while variances are 0.05, 0.1, 0.1, 0.55, 1.5, 10.5, 55, and 105 which is the case of overdispersion . Column 2 – 4 Table 4.1-4.19 is the modified Fuzzy c method of Bezdek (1981) and Yang (1993), this methodology has extensively discussed in chapter three of this study with unique membership function constructed to address the severity of overdispersion. According to Zimmermann (2014), Yang (1993), Yang (2017), Yang (2009a), Yang(2009b), "vaguess and lack of information can be successful modeled by the Fuzzy method , which is the case of overdispersion.

In line with Bezdek (1981) and Yang (1993), overdispersion has been classified into different degree of severity; namely Not severe, Moderate, Severe, and Very Severe with their corresponding membership function in Column 2 – 4 of Table 4.1-4.19. This study fills the gap in the research and in literature relating to the problem of overdispersion in count data model as presented on .Table 4.20 – 4.22 presented the point and threshold for modification. Table 4.20 – 4.24 fill the gap in research for when overdispersion should be modified unlike Famoye (2004), Trivedi and Cameron (1986), Shmeulli (2004), Consul et'al (1992) that modified model without considering when overdispersion is serious cause of concern. In this research Poisson model should be modified when dispersion percentage less than 28.24%, Negative Binomial less than 74.86%,ComPoisson less than 22.20%, and Generalised Poisson model less than 29.20%.

Table 4.23 examines for the presence of collinearity among the explanatory variables used to model the crashes of the accident data. Nwakwo (2015) opined that "correlation between the explanatory variables should be examined before fitting the model .The Variance inflation factor (VIF) will be computed, if VIF greater than 10 indicates the presence of collinearity among the variables .The explanatory variables are Speed violation (SPV), Using of Phone while driving (UPD), Overloading (OVL), Dangerous driving (DGD), and Sleeping while driving (SOS).The Variance inflation factor for Speed violation is 1.434,), Using of Phone while driving is 1.249, Overloading is 1.327. Dangerous driving is 1.229 and Sleeping while driving is 1.029 which indicates each of VIF than one , it means the variables are suitable to study the crashes of the accident data.

Figure 4.1-4.6 is exploratory data analysis to know the distribution of the data of life study .this shows from the plot that data are normally disturbuted. Table 4.26 is the Fuzzy set method for classifying the accident data. The mean of the life study is 255.10 and the variance is 52 579.23 which show a case of overdispersion Famoye (2004), modified Poisson model when mean was 0.76 and variance was 1.35, in light with Famoye (2004) the mean and variance of the accident data is 255.10 and the variance is 52 579.23; in addition dispersion percentage was computed to know the severity of the dispersion. This research fill the gap by computation of dispersion percentage from the methodology in chapter three and according to the modified method of Bezdek(1981) and Yang (1983), Zimmerman (2001), according to this study the dispersion percentage is 99.5% which classified as very severe with the membership of 1, thus this study fill the gap in research by classifying overdispersion into different category with the modified Fuzzy set method.

Table 4.27-4.30 is the parameter estimation of the models used to fit the life study.It is to examine the significance of the variables to the life study. Hypothesis was carried out at 5% level of significant. The result reveals that all the variables contributed to the incidence of Road traffic crashes , for the P-value less than the level of significance. In line with Nwakwo(2015), Famoye(2004), Famoye(2006), Srinvas(2008), Shmuelli(2005) that those variables be included when modeling Road Traffic Crashes.

Table 4.31 is the Akaike Information Criteria (AIC). This study attempt the appropriate model for the life study. The four models were used to fit data. AIC is used for model selection. From the Table 4.31, the least AIC is considered to be suitable model for the count data .The Table reveals that Generalised Poisson Model is the model with least value of AIC that is 2205.42 in line with Famoye(2004).

Figure 4.7-4.10 is the Quantile- Quantile- plot(QQ) plots for the four models to verified if the samples from the parent population, the plots depict the samples are from parent population .

The discussion of the results, therefore, have shown the gap this study has filled in research and literature.

CHAPTER FIVE

SUMMARY, CONCLUSION, AND RECOMMENDATIONS

5.1 Summary

This section is the summary of the findings discussed in the previous chapters with the conclusion and recommendations. It was revealed from the literature that there is a dearth of literature on the threshold for the point for modification of overdispersion when it occurs. Previous studies did not consider the threshold to know if there is need for modification when there is overdispersion. It was observed that researchers modify existing models when there is a slight difference between the mean and variance of the Poisson model and oftentimes, the developed models sometimes provide a poor fit when used to model count data. Many modified the model even when there may be no need for modification

Many of the scholars modified the Poisson model to capture the problem of overdispersion without first examining the cause which contributed to the problem of overdispersion and the excess zero. Oftentimes, apparent overdispersion may be encountered when dealing with count data and many of the researchers still go ahead to modify this model when the restructuring of the model will resolve this problem Hilbe(1998) but researchers just modify when there is no need for this. This is the motivation of this study in providing the threshold for the modification of the models considered in this study with this threshold proposed the researchers should cross-check the severity of the overdispersion before modifying the overdispersion model. Overdispersion was classified into different degrees of severity with the use of a Fuzzy set namely: Not severe (Mild), Moderate, Severe, and Very severe with the construction of the membership function for each of the classes and the threshold were determined by Averaging method.

The summary of the finding from the results show that overdisprsion is severe for the following models when the dispersion percentage is 28.24% for Poison model, 22.20% for Com-Poisson, 74.86% for Negative Binomial and 29.20 for Generalised at this threshold the model should be considered for modification. The result for the Road Traffic Crashes revealed that there is evidence of overdispersion inherent in the data the ratio of residual deviance to degree of freedom is 42.1. The dispersion percentage of the road crashes is 99.5% which belongs to the class of very severe with membership function of 1. The explanatory variables namely Speed violation (SPV), Using of Phone while driving (UPD), Overloading (OVL), Dangerous driving (DGD), and Sleeping while driving (SOS)all contributed to the causes of the road crashes because the p- value for SPV, UPD, OVL, DGD, SOS are $2e^{-16}$, $2e^{-16}$, $2e^{-16}$, $2e^{-16}$, and $2e^{-}16$ respectively less than 0.05 level of significance. The AIC for model selection for the four models for PO, NB, CP and GP is 8836.70, 2211.03, 8657 and 2205.42 respectively. This shows that GP has the least value and appropriate for modeling the accident data.

5.2 Conclusion

In line with the simulation study and accident data analysis, it can be reasonably concluded that the aim and objectives of this research were achieved. The simulation study provided the threshold for the modification of the models under consideration. Each threshold for the modification of the model had been determined and membership function was also constructed for the study. With the use of the Akaike information criterion, the appropriate model for the count data was selected which showed that the Generalised Poisson model was the most appropriate model for the count data.

5.3 Recommendations

As earlier mentioned this research has contributed to research that there is a need to consider threshold when overdispersion is a serious cause of concern. Based on the assumptions and analyses when there is a problem of overdispersion, there is no need to modify these models when

- Poisson model is less than 28.24%,
- Negative Binomial is less than 74.86%,
- ComPoisson is less than 22.20%, and
- Generalised Poisson model is less than 29.20%

At these points, these models should not be considered for modification when the above thresholds are not attained.

5.4 Contributions To Knowledge

Different models have been modified to solve the problem of overdispersion in count data but there is still dearth of research when overdispersion is of serious cause of concern. This research, therefore, has contributed to knowledge by:

- Categorised different percentages of overdispersion and determined when overdispersion is a serious cause of concern.
- Different membership functions have been constructed to determine the threshold when the models should be modified.
- Threshold for modification of Poisson model, Negative Binomial, Com-Poisson and Generalised Poisson models are 28.24%, 74.86%, 22.20%, 29.20% respectively.

5.5 Suggestions for Further Studies.

The study observed the possible areas of extension of this work include:-

- Determination of threshold for modification for other count data models.
- Consider Bayesian method rather than the classical method used to maximise the use of apriori information about the models.

criterion, the appropriate model for the count data was selected which showed that the Generalised Poisson model was the most appropriate model for the count data.

REFERENCES

Agarwal, D.K., Gelfand, A., Citron-Pousty, S. 2002. Zero-inflated model with application to spatial count data. *Environmental and Ecological Statistics*. 9: 341-355.

Alaba , O. O, Olubusoye, O. E and J.O. Olaomi 2017. Spatial Patterns and determinants of Fertility levels among women of childbearing age in Nigeria. *South African family Practice 59(4): 143-147.*

Alfonso, P, Jaume, V, Losila, J.M, Rafael, J. (2013)."Overdispersion in Poisson regression model. A comparative simulation study" : correction. *Methodology European Journal of Research Methods for the Behavioural Social Sciences 9(4):178.*

Bezdek, J.C 1981. Pattern recognition with Fuzzy Objectives Function algorithms. *Pienum New York* . Chapter 8:1-10.

Bihn, E.A, Fick, B. J., Pahl, D.M, Stoeckel, D. M. Woods, K.L., Wall, G.L. 2017. Geometric mean, statistical threshold values and microbial die-off Rates. *Produce Safety Alliance.GM & STV1:1-4.*

Bohning, D., Dietz, E., Schlattman, P., Mendonca, L., Kirchner, U. 1996. The zero-inflated model. *Journal of the Royal Statistical Society, Series A*. 162: 195 -209.

Bonafede , E.; Frank P. ;Stephane R.; Cinzia V. 2016. Modeling overdispersion Heterogeneity in differential expression analysis using mixtures. *Biometrics 2016 Sep. 72(3):804-814.*

Cameron, A.C., Trivedi, P.K. 1986. Econometric models based on count data: comparisons and applications of some estimators and tests. *Journal of Applied Econometrics*. 1: 29-53.

Cameron, A.C., Trivedi, P.K. 1998. Regression Analysis of Count Data. *New York: Cambridge University Press.* Chapter 3:67-89.

Castillo (1998). Weighted Poisson Distribution for Over-dispersion and underdispersion. Situation. *Annals of Institute of Statistical Mathematics* 50:567-585.

Consul, P.C., Famoye, F. 1992. Generalized Poisson regression model. *Communications in Statistics (Theory & Method)*. 2(1): 89-109.

Consul, P.C., Jain, G.C. 1973. A generalization of the Poisson distribution. *Technometrics*. 15:*791-799.*

Consul, P. C. 1989. Generalized Poisson and Negative Binomial *Technometrics*. 10: 917-920.

Conway, R. W., Maxwell, W. L. 1962. A queuing model with state dependent service rates. *Journal of Industrial Enginweering.* 12: 132–136.

Durmus and Guneri 2020. An application of the Generalised Poisson model for overdispersion for overdisperion number before strikes between 1984 and 2017. *The Journal of Operations Research , Statistics, Econometrics and Management information systems.*
Volume 8, issue 2, 2020.

Efron 1986. Double Exponential families and their uses in Linear Generalised Regression. *Technical Report. No 239:39-45.*

Endo, A 2020. Estimating the overdispersion in COVID-19 transmission using outbreak seizure outside China.*www.ncbi.nlm.nih.gov Vol 6(9):1884-1893.*

Famoye, F., Wulu, J.T., Singh, K.P.2004. On the generalized Poisson regression model with an application to accident data. *Journal of Data Science*. 2: 287-295.

Famoye, F., Singh, K.P. 2006. Zero-inflated generalized Poisson regression model with an Application to domestic violence data. *Journal of Data Science*. 4: 117-130.

Gardner, W. ,Mulvey E. P., Shaw C. E 1995. Regression analyses of count and rate: Poisson, Over-dispersed Poisson and negative binomial models. *Psychological Bulletin* 118: 392–404.

German, R. 2000. Poisson Models for count data. *Online Lectures on Poisson model* . Chapter 4:1-14

German, R. 2013. Models for count data with over-dispersion. *Online Lectures* Chapter 1:1-7

Greene, W. 1994: Accounting for excess zeros and sample selections in Poisson regression and Negative Binomial regression models. New York University Working Paper No. Ec-94-10. Chapter 1:1-37

Greene, W. 2008. Functional forms for the negative binomial model for count data. *Economics Letters. 99: 585-590.*

Gschlobl, S. 2013. Modeling count data with over-dispersion and spatial effects. *Statistical Paper 49(3):1-12*

Guikema, S. D and Coffelt; J.P. 2008. A flexible count data regression model for Risk Analysis. *Risk Analysis 32 :213-223.*

Gul Inan 2017. Comments on “Marginalised Multi-level hurdle and zero-inflated model for overdispersed correlated count data with excess zeros”. *Statistics in Medicine, Vol 3, 2.*

Hansen, B.E 2000. Sample splitting and Threshold estimation. *Econometrica, Vol. 68; No , 575-603.*

Harrison, X. A. 2014. Using observation level random effects to model over-dispersion in count data in Ecology and Evolution. *Peer J2(1):e616.*

Hilbe, J. 2000. Negative Binomial Regression. *Cambridge, UK: Cambridge University Press*. Chapter 8:21-50

Hoissein Arsham and Miodrag Lovric 2013. Bartlett's test. https://www. researchgate.net*//publication/252322443.Researchgate Vol 5(6:80-85.*

Ismail, N, and Abdul A. J, 2007. Handling overdispersion with Negative Binomial and Generalised Poisson regression. *Casualty Actuarial Society E Forum Spring 2007.Page 103-156*

Ismail, N and Hussein, Zamani 2013. Estimation of claim count data using Negative Binomial, Generalized Poisson, Zero inflated Negative Binomial and Zero-inflated generalized Poisson regression models. *Casualty Actuarial Society E Forum Spring 2013.Page 1-30.*

Joe and Zhu 2005. Generalized Poisson distribution, the property of mixture of Poisson Distribution and negative binomial distribution. *Biometrical Journal 47:219–229.*

John Hinde and Demetrio G.B 1998. Over-dispersion: models and Estimation. Journal of Computational Statistics and Data Analysis. *Elsevier*, Vol. 27(2),pages 151-170.

Jonathan Marchini 2008. Poisson Distribution. *www.Stats.ox.ac.uk*. Chapter 3:1-8

Kimberly, F.S. and Shmueli, G. 2010. A flexible regression for count data. *The Annals of Applied Statistics Vol. 4 No 2,943-961*

Kimberly, F. S., Shared, B. and Shmueli, G. 2012. The Com-Poisson model for count data: a survey of methods and application. *Applied Stochastics Models in Business and Industry (2012):28:104-116.*

King, G 1988. Statistical models for Political Science events. Bias in conventional procedures and evidence for the exponential Poisson regression model. *Mathematical Population Studies, Vol. 32, 838-863.*

King, G 1989. Variance specification in event count models: from restrictive to generalized estimator. *American Journal Of Political Science* 33,762-784

Kwang, H. Lee 2005. First Course on Fuzzy Theory and Application. *Springer Berlin Heidelberg New York.*. Chapter 1-51

Krull, M. 2020. Comparing Statistical analysis to estimate threshold in ecotoxicology hppt://*www.ncbi.nlm.nih.gov.*Vol15(4):1-16

Lambert D.1992. Zero Inflated Poisson Regression with an application to Defects in Manufacturing. *Technometrics* 34, 1-14.

Lawal, A. S., Buliaminu, K and Peter, Oke 2012. Modeling Fatalities of Road accidents in Nigeria. *Research in Logistic and Production* https://www.researchgate.net/ *publication/304676305.Researchgate.*Vol 2, N0 3, pp 259-271.

Lawless, J.F. 1987. Negative Binomial and mixed Poisson Regression. *Canadian Journal of Statistics*. 15(3): 209-225.

Lee, J.H, Han, G., Fulp W. J. and Giuliano A. R. 2012. Analysis of overdispersed count data :application to Human Papillomaurus Infection in Men (HIM) study. *National Institute of Health, 2012 June 140(6),1087-1094.*

Lindsey, J.K. 1999. On the use of corrections for overdispersion. *Journal of Royal Statistics Society Series C (Applied Statistics) 48*, 553-561.

Lord, D. and Ivan J.N. 2005. Poisson, Poisson gamma and zero inflated model. *The Annals of Applied Statistics. Vol 37, Issue 1. 2005.pp 35-46.*

Lord, D.2008. Application of the Conway-Maxwell Poisson generalized linear model for analysing motor vehicle crashes. *Accident Analysis and Prevention 40,1123-1134.*

Lord, D., Geedipally, S.R., Guikema, S.D. 2010. Extension of the application of Conway- Maxwell-Poisson models: analyzing traffic crash data exhibiting under-dispersion. *Risk Analysis*. 30(8): 1268-1276.

McCullagh, P., Nelder, J.A. 1989. Generalised Linear Models (2nd Edition). *Chapman and Hall: London.*Chapter 5:70-80

Mei-Chen, H., Pavlicova, M. and Nunes, V. E. 2017. Zero inflated and Hurdle Models of count data with extra zeros: example treat sample from an HIV-Risk reduction intention. *National Institute of Health Public Access 3(5):67-375.*

Mironov 2006. Threshold selection using the rank statistics. *Paper work Frontliers Public Health, Vol. 5 No 2, 130-134.*

Javidi, M. M., and Mansoury S 2017. Diagnosis of disease using an ant colony gene selection method based on information gain ratio using fuzzy roughs sets. *Journal of Particle Science and Technology , (2017) 175-186.*

Neelon, B.H., O'Malley, A.J., Normand, S.T.2010. A Bayesian model for repeated measures zero-inflated count negative binomial distribution. *Biometrical Journal*. 47: 219-229.

Nelder, J.A and R.W. Wedderburn 1972. Generalised Linear Models. *Journal of the Royal Statistical Society. 135(3):370-384.*

Nwankwo, Chike H and Nwaigwe Godwin I.2015. Statistical Model of Road Traffic Crashes Data in Anambra State, Nigeria: A Poisson Regression Approach. *International Journal of Scientific & Technology Research Volume 4, Issue 09, September 2015.pp 226-233*

Ozmen, I., Famoye, F. 2007. Count regression models with an application to zoological data zeros. *Journal of Data Science*. 5: 491-502.

Palmer, A., Jau Containing structural me, V., Losila, J. M. and Rafael, Jimenez 2013. Overdispersion in the Poisson regression model: A comparative simulation correction. *Researchgate.*

Payne, E. H., James. W. H., Leonard E. E, Viswanathan, R., Anbesaw S. and Malugeta G. 2015. Statistical Methods in Medical Research 0(0)1-28.

Renshaw, A.E.,1994. Modeling the claims process in the presence of covariates. *ASTIN Bulletin*. 24(2): 265-285.

Richard, Berk and Macdonald J. 2007. Over-disperse and Poisson regression. *Journal of Qualitative Criminology* 24(3):269-284.

Ridout, M., Demetrio, C. G. B. and Hinde, J. (1998). Models for count data with many zeros. *Biometrics Society* 179-190.

Ridout, M.S., Hinde, J.P., Demetrio, C.G.B. 2001. A score test for testing a zero-inflated Poisson regression model against zero-inflated negative binomial alternatives. *Biometrics*. 57: 219-223.

Shmueli, G., Minka T., Kadane, J.B., Borle, S., Boatwright, P.B. 2005. A useful distribution for fitting discrete data: revival of the Conway–Maxwell–Poisson distribution. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 54: 127–142.

Srinivas, R. Geedipally 2008. Examining the application of Conway-Maxwell Poisson Models for Analyzing Traffic crash data. Researchgate.Vol 6,No14 pp 1-12

Stram, D.O., Lee, J.W. 1994. Variance components testing in the longitudinal mixed effects model. *Biometrics*. 50:1171-1177.

Stram, D.O., Lee, J.W. 1995. Correction to "Variance components testing in the longitudinal mixed effects model". *Biometrics*. 51: 1196.

Syeda B. Z. and Talmeez Hussain 2017. Human Blood Pressure and Body

Temperature Analysis using Fuzzy logic control system. *International Journal of Computer Science and Network Security, Vol 17, pp 12.*

Theiwall M. 2005. The Precision of the arithmetic mean, geometric mean and Percentiles for Citation data: An experimental simulation modeling approach. *Journal of Informetrics 10(1):1-21*

Ulm K.1991. A statistical method for assessing a threshold in epidemiological studies. *Statistics in Medicine,Vol.10.341-349.*

Ver Hoer Hay M. 2007. Quasi-Poisson vs negative binomial regression: how should we model over-dispersed count data? *Ecological Society of America. 88(11), 2007, pp.2766–2772.*

Vuong, Quang H. 1989. *Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica. 57(2):307-333*

Wang, W., Famoye, F. 1999. Modeling household fertility decisions with generalised Poisson regression. *Journal of Population Economics*. 10: 273-283.

Wei C., Yi L., Xiao, Longfei Xu (2020): Integrative analysis of spatial heterogeneity and overdispersion of crime with a Geographically weighted Negative Binomial. *International Journal Geo- information. Vol 9 p.60*

Winkermann and Zimmermann C. 1994. Count data model for demographical data. *Journal of Mathematics Population Studies* 4, 205-221.

Winkelmann, R. 2000. Econometric Analysis of Count Data. *Heidelberg: Springer Verlag.Chapter 4:201-225*

Witold Pawlus, Hamid K. Kjeil R. 2012. A Fuzzy logic approach to modeling a vehicle crash test. *Central European Journal of Engineering Vol 3,67-79.*

Yan, Yuan., Wanhua, S. and Mu Z. 2015. Threshold free measure for assessing the performance of medical screening tests. *Frontiers in Public Health,3,p.57.*

Yang, Z.(1993): Fuzzy Class Maximum Likelihood (FCML).*Science Direct* .Vol 57, issue 3, page 365-375

Yang, Z., Hardin, J.W., Addy, C.L. Vuong, Q.H. 2007. Testing approaches for overdispersion in Poisson regression versus the generalized Poisson model. *Biometrical Journal*. 49: 565-584.

Yang, Z., Hardin, J.W., Addy, C.L. 2009a. A score test for overdispersion in Poisson Regression based on the generalized Poisson-2 model. *Journal of Statistical Planning and Inference. 139: 1514-1521.*

Yang, Z., Hardin, J.W., Addy, C.L. 2009b. Testing overdispersion in the zero-inflated Poisson model. *Journal of Statistical Planning and Inference*. 139: 3340-3353.

Yaotian, Zou and Dominique, L. 2013. Evaluating the Double Poisson Generalized Linear Model approximate normalizing constant with high accuracy. *Accident Analysis & Prevention Vol 59. 2013, page 497 – 505*

Yoonseok 2020. Threshold Regression Model. *Centre for Policy Research-Spring 2020.*25 :278-295.

Zadeh, L.A 1965. Fuzzy set. *Information and Contro.,Vol.8(1965),pp 338-353*

Zhou, Qi 2018. Empirical approach to threshold determination from the delineation of built area with road network road network data. *National Centre for Biotechnology 13(3)2018:1-25.*

Zimmermann, H.J. 2001. Fuzzy set Theory and its Application. *Text Book Third Edition.*Chapter 1:1-20

APPENDIX

TABLE 1: DATA PRESENTATION FOR THE ACCIDENT DATA

Crashes	SPV	UPD	OVL	DGD	SOS
87	27	0	0	9	0
288	65	0	0	50	0
57	31	0	1	8	0
272	115	2	2	29	0
493	114	0	5	27	9
95	4	1	0	5	0
290	143	2	14	53	2
14	5	0	0	2	0
197	81	2	5	51	0
290	152	1	1	664	1
248	93	1	3	39	4
239	137	0	0	21	1
60	36	0	1	7	0
263	111	1	2	31	1
1395	578	3	6	118	4
179	35	0	0	34	1
241	64	0	0	37	0
101	32	0	1	9	0
525	196	3	7	60	1
404	122	0	1	34	0
159	57	3	18	18	0
149	69	0	4	10	0
254	30	0	0	8	0
199	4	0	1	27	2
321	78	0	4	43	0
878	198	4	13	180	5
602	192	3	1	79	4
298	112	2	7	27	0
285	152	0	2	63	7
266	140	0	0	19	1
272	122	1	8	21	0
245	47	0	0	23	1
137	48	0	0	9	1
158	42	2	3	1	1
83	15	0	0	29	0
70	8	0	0	19	2
266	41	1	4	60	0
86	49	0	1	4	0
163	31	0	1	18	1

50	25	0	1	10	0
255	96	2	3	40	1
226	61	0	4	15	2
69	20	1	0	1	0
388	128	5	4	49	8
9	4	0	0	1	0
138	49	0	3	21	1
200	94	2	0	30	1
287	100	3	3	45	0
274	139	0	1	21	0
55	29	0	0	4	0
250	26	0	5	18	1
1342	393	7	3	135	3
189	68	1	3	17	0
170	40	0	0	36	0
149	70	2	0	10	2
502	147	0	6	91	1
269	155	2	1	47	0
225	50	6	5	12	0
143	65	0	1	7	3
331	73	0	1	9	0
207	22	0	0	15	4
403	84	0	6	52	8
798	215	2	5	128	3
523	197	1	7	49	3
428	218	0	3	26	0
211	85	0	0	41	3
266	110	0	4	34	1
270	147	0	2	41	2
236	36	0	1	24	1
96	40	0	1	5	0
164	52	3	4	6	1
106	9	0	0	38	3
53	14	0	0	13	0
203	54	1	3	24	2
102	67	0	2	3	0
162	60	0	1	16	0
51	28	1	0	5	0
230	71	3	6	28	5
308	88	0	2	20	2
45	15	0	1	1	4
342	191	1	1	33	0
26	6	0	0	2	0
98	48	1	0	0	0
129	62	0	0	5	3

269	71	0	3	30	5
268	135	0	6	13	0
49	23	1	1	4	3
238	59	0	2	4	0
1373	483	2	6	194	3
18	85	1	1	7	0
119	51	1	0	10	1
144	69	2	7	13	1
715	335	0	8	115	1
390	206	0	4	31	1
247	60	4	6	16	1
145	67	2	1	15	0
27	117	0	0	7	1
197	62	0	3	15	5
441	145	1	3	45	4
530	153	3	9	40	8
535	298	4	2	35	3
387	229	0	5	3	0
255	108	0	2	49	4
245	88	0	2	23	4
321	152	1	2	32	5
244	43	0	3	18	0
111	51	0	0	5	0
122	37	2	5	18	6
156	17	0	0	45	1
101	16	1	2	17	0
164	52	1	3	8	1
68	42	0	1	5	1
129	63	0	2	2	0
90	51	0	0	11	0
259	115	1	2	35	2
338	105	1	4	30	9
45	22	1	0	0	1
324	237	0	0	20	0
36	15	0	0	2	0
64	44	0	0	12	0
134	65	0	3	3	0
277	59	0	0	19	1
226	117	1	4	15	0
56	32	0	0	10	0
195	99	1	2	14	1
1106	623	1	7	133	1
159	80	2	0	22	0
156	94	0	0	22	1
218	109	0	3	16	0

755	480	3	8	92	0
303	206	4	5	12	0
250	117	5	1	24	0
153	81	0	4	9	3
374	204	0	1	4	0
183	98	0	3	10	0
425	217	1	3	42	1
502	237	2	5	70	3
516	254	1	10	26	3
378	199	0	2	15	1
357	205	3	1	48	6
213	138	0	1	10	3
346	186	2	5	48	2
235	82	1	1	1	0
102	50	3	0	10	1
91	31	0	1	13	2
137	15	2	0	65	20
82	22	0	3	8	0
101	46	2	2	2	0
68	26	1	2	11	2
889	491	2	1	133	0
472	51	0	1	3	0
262	50	0	0	3	0
115	54	0	0	19	0
244	85	0	3	20	6
40	13	0	0	0	0
140	135	0	0	0	0
31	15	0	1	0	0
53	26	0	1	6	0
95	46	0	3	6	0
67	29	0	0	9	4
151	89	0	0	4	0
62	24	0	0	5	0
132	64	2	2	8	0
141	64	0	3	31	0
97	35	0	0	5	4
164	82	1	11	3	0
701	362	2	12	58	0
252	150	2	11	10	4
217	108	2	5	21	0
94	73	0	1	5	0
275	161	0	3	7	0
190	87	0	0	11	0
265	141	0	1	14	2
730	389	0	4	32	2

350	197	2	4	16	9
466	240	0	4	24	3
398	224	0	2	14	4
236	110	0	1	9	2
355	175	0	2	47	4
187	67	3	1	7	0
88	38	2	1	3	1
72	29	1	5	10	0
103	14	0	3	42	14
97	37	0	2	17	2
94	38	0	1	11	5

R Package source

```
getwd()

setwd("C:/Users/ACER/Desktop/analysis")


n=20

# Specify the number of replications

# number of retained replications

s = 20



# store all draws in the following matrices

# initialize them here

m = matrix(nrow = n, ncol = 1)

v = matrix(nrow = n, ncol = 1)

d = matrix(nrow = n, ncol = 1)

g = matrix(nrow = n, ncol = 1)

M = matrix(nrow = n, ncol = 1)


################### Code for Dispersion########################################

Simulation for Poisson Distribution

for (i in 1:n){

Y1 <- rpois(s, 2 * exp(rnorm(s, mean=0.01, sd=0.05)))

m[i,]<-mean(Y1)

v[i,]<-sd(Y1)^2

d[i,]<-abs((mean(Y1)-(sd(Y1)^2))/(sd(Y1)^2))*100

}
```

```
#Simulation for Negative Binomial distribution
#for (i in 1:n){
#x1 <- rnbinom(s, mu = 4, size = 1)
#m[i,]<-mean(x1)
#v[i,]<-sd(x1)^2
#d[i,]<-abs((mean(x1)-(sd(x1)^2))/(sd(x1)^2))*100
#}


#Simulation for Conway Maxwell
#install.packages("degreenet", dependencies=T)
#library(degreenet)
#for (i in 1:n){
#x1 <- simcmp(n=5000, v=c(7,3))
#m[i,]<-mean(x1)
#v[i,]<-sd(x1)^2
#d[i,]<-abs((mean(x1)-(sd(x1)^2))/(sd(x1)^2))*100
#}
#testOverdispersion(d)


p=10
r=40
w=70
Group = g
Member = m
```

```
result = data.frame(m,v,d,g,M)

colnames(result)<- c("mean","Variance","Dispersion","Group","Member")

head(result)



result$Group[result$Dispersion <= 10] = "Not Severe"

result$Group[result$Dispersion >= 10 & result$Dispersion <= 40]="Moderate"

result$Group[result$Dispersion >= 40 & result$Dispersion <= 70]="Severe"

result$Group[result$Dispersion >= 70]="Very Severe"

head(result,15)



d1=result$Dispersion[result$Group == "Not Severe"]

d2=result$Dispersion[result$Group == "Moderate"]

d3=result$Dispersion[result$Group == "Severe"]

d4=result$Dispersion[result$Group == "Very Severe"]



p=10
```

```
r=40
w=70

M1= rep(0,length(d1))
M2=(4*d2-r)/210
M3 =d3/70
M4=rep(1,length(d4))



result$Member[result$Dispersion <= 10] = "0"
result$Member[result$Dispersion >= 10 & result$Dispersion <= 40]=M2
result$Member[result$Dispersion >= 40 & result$Dispersion <= 70]=M3
result$Member[result$Dispersion >= 70]="1"
result
result$Member
w = table(result$Member)
t = as.data.frame(w)
names(t)[1] = 'membership'

as.data.frame.table(table(membership = result$Member))
result20=head(result,n)
result20
group=result20[,4]
table(group)
write.csv(result20,"rpois20.csv")
```

```
#hist(Y1<-(rpois(n, mean=0.01)), breaks=seq(0, 30), col="grey60", freq=F, ylim=c(0, 0.45), las=1, main="", xlab="Y")

#Membership Category
result$Dispersion

library(RNGforGPD)
GenUniGpois(5, -0.4, 100, method = "Inversion")
GenUniGpois(2, 0.9, 100, method = "Branching")
GenUniGpois(12, 0.5, 100, method = "Normal-Approximation")

data <- GenUniGpois(3, 0.9, 100, method = "Inversion", details = FALSE)

data <- GenUniGpois(10, 0.4, 10, method = "Chop-Down", details = FALSE)

n=100
# Specify the number of replications
# number of retained replications
s = 20

# store all draws in the following matrices
# initialize them here
m = matrix(nrow = n, ncol = 1)
v = matrix(nrow = n, ncol = 1)
d = matrix(nrow = n, ncol = 1)
```

```
g = matrix(nrow = n, ncol = 1)

M = matrix(nrow = n, ncol = 1)



#Simulation for Generalized Poisson distribution

#for (i in 1:n){

#data <- GenUniGpois(3, 0.9, 100, method = "Inversion", details = FALSE)

#x1 <- data$data

#m[i,]<-mean(x1)

#v[i,]<-sd(x1)^2

#d[i,]<-abs((mean(x1)-(sd(x1)^2))/(sd(x1)^2))*100

#}
```