ABSTRAK

Disertasi ini mencadangkan satu taburan baru untuk pemodelan data serakan terkurang, sama dan terlebih. Taburan in muncul sebagai kes istimewa daripada perjalanan rawak yang diubah pada satah dan taburan ini termasuk binomial, binomial negatif dan shifted binomial negatif sebagai kes-kes khasnya. Ciri-ciri, ujian hipotesis untuk serakan sama, kajian simulasi kuasa, penganggar kebolehjadian maksimum (PKM) dan kaedah jarak kuasa dua berdasarkan fungsi penjana kebarangkalian dipertimbangkan. Selain itu, taburan yang dicadangkan ini dibandingkan dengan taburan yang sedia wujud seperti taburan COM-Poisson dan taburan Generalized Poisson dalam pemodelan serakan. Taburan ini telah dibuktikan sebagai model yang fleksibel melalui aplikasi kebagusan penyuaian kepada empat set data nyata. Tambahan pula, dua taburan bivariat baru diterbitkan berdasarkan taburan yang telah dicadangkan sebagai taburan alternatif bivariat diskrit dengan menggunakan konvolusi daripada dua taburan bivariat dan kaedah klasik penurunan trivariat. Didapati, taburan bivariat baru ini mengizinkan lebih banyak fleksibiliti dalam pemodelan dan had kepada korelasi antara dua pembolehubah rawak adalah kurang. Ciri-ciri dan anggaran parameter taburan bivariat baru ini telah disediakan. Selanjutnya, inferensi statistik untuk taburan COM-Poisson dan isu-isu pengiraan turut diselidik. Ujian untuk serakan sama, kajian tentang kekuatan statistik dan anggaran parameter dengan menggunakan kaedah PKM dipelajari. Disertasi ini juga mempertimbangkan satu sukatan jarak berdasarkan fungsi penjana kebarangkalian bagi anggran parameter. Prestasi dan keteguhan bagi statistik yang dicadangkan dalam anggaran parameter diselidik di bawah taburan binomial negatif melalui cara simulasi Monte Carlo, khususnya dibandingkan dengan kaedah PKM dan jarak Hellinger minimum.

ABSTRACT

This thesis proposed a new distribution to model under-, equi- and over-dispersion in count data. It arises as a particular case of a modified random walk on the plane, and includes the binomial, negative binomial and shifted negative binomial as special cases. Some properties, test of hypothesis for equi-dispersion, simulation study of power, parameter estimation by maximum likelihood and a squared distance method based on the probability generating function are considered. The proposed distribution is compared with existing distributions like the COM-Poisson and generalized Poisson for modelling dispersion. It is shown to be a flexible model in applications by illustrating its goodness-of-fit to four real data sets. In addition, two new bivariate distributions are derived from proposed distribution as the alternative bivariate discrete distributions by applying the convolution of two bivariate distributions and the classical trivariate reduction method. It is found that the new bivariate distributions permit more flexibility in modelling and less limitation on the correlation between the two random variables. The characteristic and the estimation of the parameters of the new bivariate distributions are provided. Furthermore, the statistical inference for the COM-Poisson distribution and computational issues are studied. Test for equi-dispersion, study of statistical power and parameter estimation by maximum likelihood are developed. This thesis also considers a probability generating function-based divergence statistic for parameter estimation. The performance and robustness of the proposed statistic in parameter estimation is studied for the negative binomial distribution by Monte Carlo simulation, especially in comparison with maximum likelihood and minimum Hellinger distance estimation.

ACKNOWLEDGEMENTS

I am heartily thankful to my supervisor, Prof. Dr. Ong Seng Huat, whose encouragement, advice, guidance and support from the initial to the final stages of this thesis, has inspired and enriched my growth as a student and a researcher. Thank you for triggering and nourishing my intellectual maturity that will benefit me for a long time to come. I am extraordinarily fortunate to have him as my supervisor.

I would like to thank my thesis examiners who have spent their precious time to read this thesis and gave crucial comments to improve it.

Thanks are also due to my family members, especially my parents who raised me with care, love and have always supported me in all my pursuits.

My special thanks go to Yap Soon Lee for his invaluable help throughout my thesis work and for keeping me organized.

Lastly, I offer my regards and blessings to those who have supported me in any way for the successful realization of this thesis.

TABLE OF CONTENTS

GLOSSARY OF ABBREVIATIONS

χ^2	Chi-square
$f(x_1, x_2)$	Joint probability mass function
$\varphi(t_1, t_2)$	Joint probability generating function
$\rho(X_1, X_2)$	Correlation between the random variables X_1 and X_2
$Cov(X_1, X_2)$	Covariance between the random variables X_1 and X_2
$Var(X_i)$	Variance for the random variables X_i
$E(X_i)$	Mean for the random variables X_i
${}_2F_1$	Gauss hypergeometric function
$\mu'_{[r]}$	*r*th descending factorial moment
N	Sample size
$\varphi_{X_1}(t \mid x_2)$	The pgf of the conditional distribution of X_1 given $X_2 = x_2$
B	Binomial
BGITD	Bivariate $\mathrm{GIT}_{3,1}$ distribution
BGPD	Bivariate generalized Poisson distribution
BNBD	Bivariate negative binomial distribution
COM-Poisson	Conway-Maxwell-Poisson
df	Degree of freedom
epgf	Empirical probability generating function
GIT	Generalization of the shifted inverse trinomial
$\mathrm{GIT}_{3,1}$	Subclass of the generalized shifted inverse trinomial distribution
gof	goodness-of-fit

GPD	generalized Poisson distribution
HD	Hellinger distance
ID	Index of dispersion
IT	Inverse trinomial
KL	Kullback-Leibler
LR	Likelihood ratio
LRT	Likelihood ratio test
MHD	Minimum Hellinger distance
MHDE	Minimum Hellinger distance estimation
MLE	Maximum likelihood estimation
MSE	Mean square error
NTA	Neyman Type A
NB	Negative binomial
pf	Probability function
pgf	Probability generating function
PIG	Poisson inverse Gaussian
pmf	Probability mass function
POI	Poisson
RS	Rao's score
SA	Simulated annealing

LIST OF FIGURES

LIST OF ALGORITHMS

LIST OF TABLES

LIST OF APPENDICES