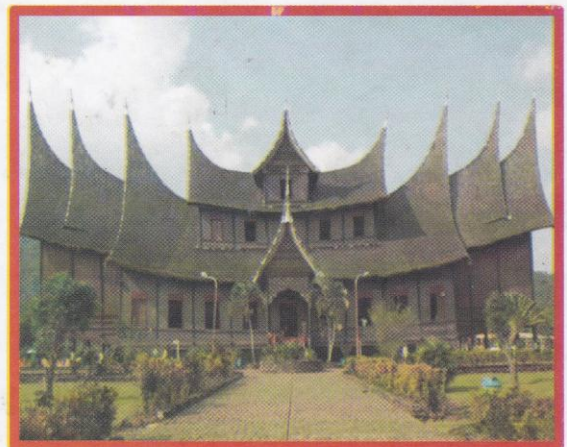
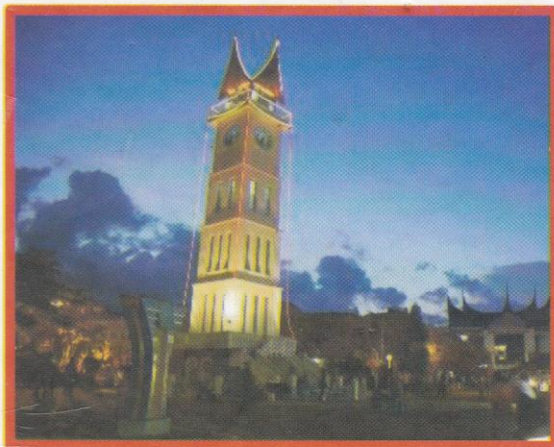


The 5th IMT - GT International Conference on Mathematics, Statistics and Their Applications **ICMSA 2009**

Editors :
I Made Arnawa, Muhafzan,
Maiyastri, Susila Bahri



June 9-11, 2009
The Hills Hotel
Bukittinggi, Indonesia

Organized by :



Proceedings of

The 5th IMT-GT

International Conference on Mathematics,
Statistics, and their Applications 2009
(ICMSA 2009)

**"Mathematics and Statistics for Industry and
Community Development"**

June 09 – 11, 2009

The Hills Hotel, Bukittinggi - Indonesia

Organized by :
Department of Mathematics, Andalas University,
Padang – Indonesia

Published by :

Department of Mathematics, Andalas University
Kampus UNAND Limau Manis
Padang 25163, Sumatera Barat,
INDONESIA.

ISBN 978-602-95343-0-6

© Copyright reserved.

The organizing Committee is not responsible for any errors or views expressed in the papers as these are responsibility of the individual authors.

June 2009

Preface

First of all, I would like to say welcome to Bukittinggi, Indonesia to all of you. It is an honour for us to host this conference. We are very happy and proud because the participants of this conference come from many countries; we have participants from Libya, Japan, Qatar, India, Malaysia, Singapore, Thailand, Iran, and many more.

Ladies and gentlemen, according to constructivism theory, mathematics comes out as a result of social construction; that's why, the outcomes of our researches in mathematics, like theorem or formula of mathematics, should be communicated in a scientific forum such as seminar or conference. Through this kind of seminar or conference, we could refine the existing theorems or we could get new ideas to produce a new one. Seminar or conference which is held annually enables us to continually develop the science of mathematics until the end of the time.

That's way, in this two-day conference, we are going to discuss around 250 papers coming from diverse aspects of mathematics ranging from analysis, applied mathematics, statistics, algebra, Computational Mathematics, mathematics education, and other related topics.

For all of us here, I would like to convey my endless appreciation and gratitude for your participation in this conference.

Thank you very much



Dr. I Made Arnawa
Chairman of the Conference

Message from Rector Andalas University

It gives me great pleasure to extend my sincere and warm welcome to the participants of the 5th International Conference on Mathematics Statistics and Application (The IMT GT's 5th ICMSA 2009) - A Joint Scientific Program organized by Universities over Indonesia, Malaysia and Thailand Growth Triangle Region. On behalf of Andalas University, let me welcome all of you to the conference in Bukittinggi, West Sumatra Province, the land of Minang kabau.

We believe that from this scientific meeting, all of participants will have time to discuss and exchange ideas, findings, creating new networking as well as strengthen the existing collaboration in the respective fields of expertise. In the century in which the information is spreading in a tremendous speed and globalization is a trend, Andalas University must prepare for the tough competition that lay a head. One way to succeed is by initiating and developing collaborative work with many partners from all over the world. Through the collaboration in this conference we can improve the quality of our researches as well as teaching and learning process in mathematics and to achieve standards and requirements applied in many developed countries. I strongly believe that this conference is and extraordinary testimony to our capacity building at international, regional and local level.

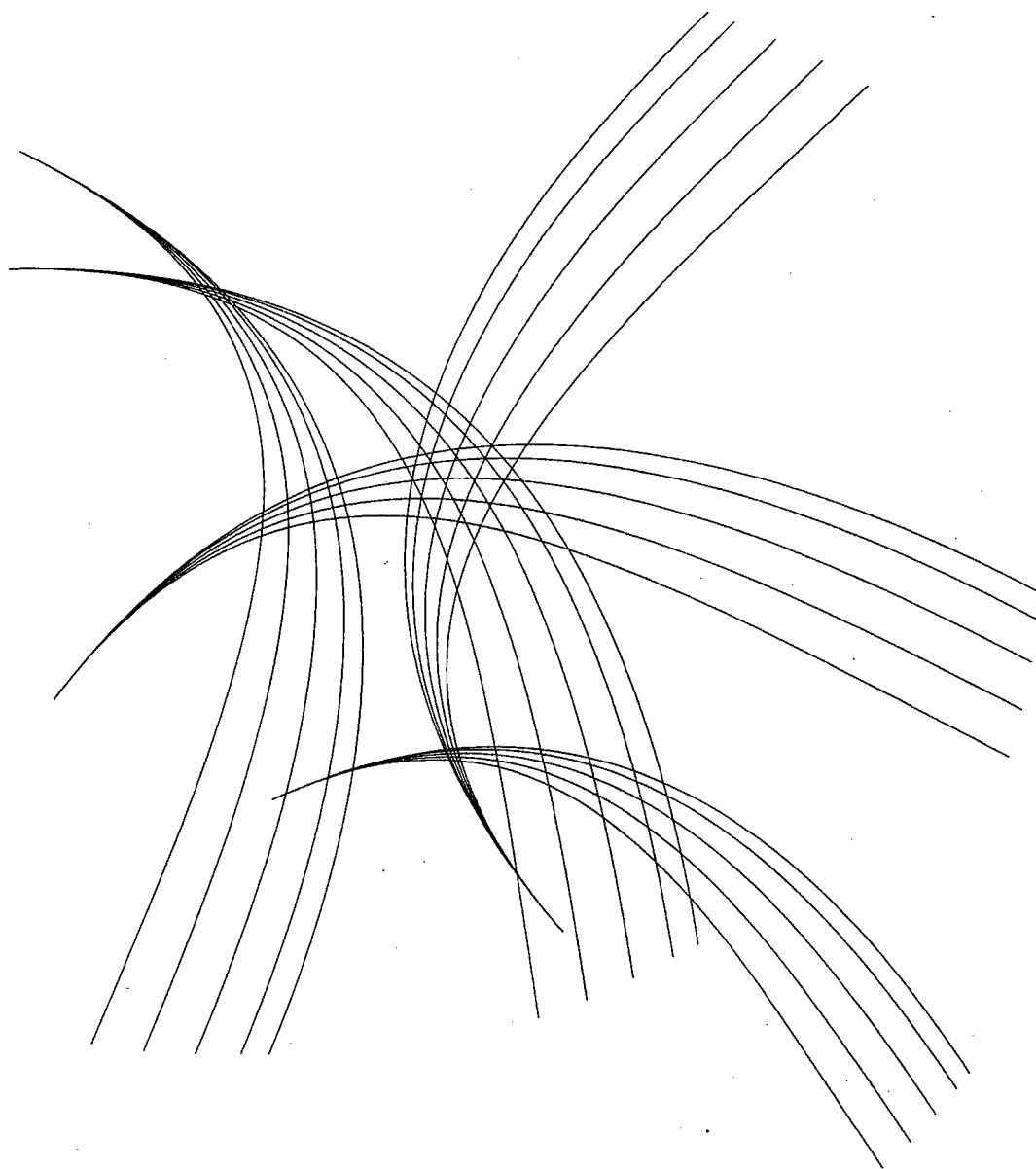
I would like to express my deep gratitude to International Scientific Committee of who has honored the Mathematics Department, Faculty of Mathematics and Natural Sciences, Andalas University to host this prestigious conference. This is a very special opportunity for us to be involved in a regional community of knowledgeable scientist in the field of mathematics, statistics and their applications. I would also like to extend my gratitude to keynote speakers, participants, and organizer of this conference for their contribution to this event. My special thank is also rendered to the local government of West Sumatra for various supports and facilities.

Finally I wish all participants a fruitful deliberation at the conference. I also wish all participants and accompanying spouses a pleasant and enjoyable stay in Bukittinggi City, West Sumatra.



Prof. Dr. Ir. Musliar Kasim, MS
Rector

Organizing Committee



ORGANIZING COMMITTEE

Patrons : Rector Andalas University
Vice Rector of Academic Affair Andalas University

Steering : Dean Faculty of Sciences Andalas university
Vice Dean of Academic Affairs Andalas University

International Scientific Committee

Prof. Dr. Novesar Jamarun (Andalas University – Indonesia)
Prof. Dr. Herman Mawengkang (University of Sumatera Utara – Indonesia)
Prof. Dr. Dato Rosihan Ali (University Sains Malaysia)
Assoc. Prof. Dr. Ahmad Izani Md. Ismail (University Sains Malaysia)
Assoc. Prof. Dr. Putipong Bookkamana (Chiang Mai university – Thailand)
Assoc. Prof. Dr. Anton Abdulbasah Kamil (University Sains Malaysia)
Assoc. Prof. Dr. Pachitjanut Siripanitch (School of Applied Statistics – Thailand)
Dr. Saib Suwilo (University of Sumatera Utara – Indonesia)
Dr. Hizir Sofyan (Syiah Kuala University – Indonesia)
Dr. Tarmizi Usman (Syiah Kuala University – Indonesia)

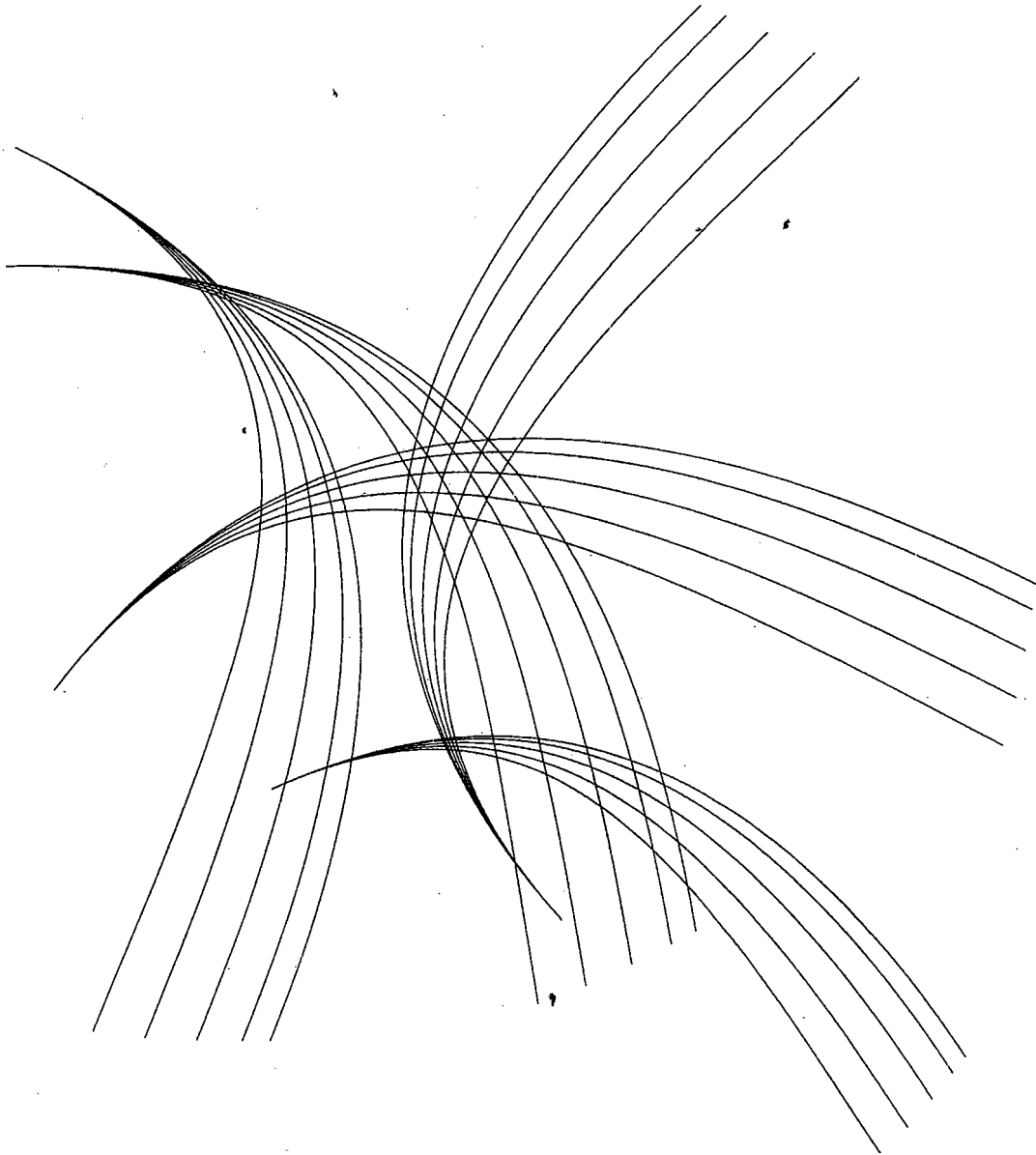
Local Scientific Committee

Dr. Maiyastri (Andalas University – Indonesia)
Dr. Susila Bahri (Andalas University – Indonesia)
Dr. Muhafzan (Andalas University – Indonesia)
Adek Tasri, Ph.D (Andalas University – Indonesia)

Committee

Chairman	: Dr. I Made Arnawa.
Secretary	: Ir. Yudiantri Asdi, M.Sc. : Izzati Rahmi, H. G., M.Si.
Transportation/Accommodation	: Drs. Syafruddin, Zulakmal, M.Si.
Treasurer	: Ir. Werman Kasoep, M.Kom., Jenizon, Msi, Ir. Hazmira Yozza, M.Si., Ir. Insanul Kamil, MS., Drs. Zulkarnain Khaidir, MS., Drs. Mulyadi, MS.
Technical Programme	: Dr. Muhafzan, Dr. Maiyastri, M.Si.
Secretariat / Documentation	: Budi Rudianto, M.Si., Narwen, M.Si., Arrival Rince Putri, M.T., Budi Rahmadya, S.Kom, Mohammad Hafiz Hersyah, S.Kom., Rahmi Eka Putri, S.Kom., Ratna Aisuwarya, S.Kom, Fajar Wisga, Berkah Fajar.

Contents



CONTENTS

Preface	i
Message from Rector of Andalas University	ii
Organizing Committee	iii
Algebra	1 - 54
A subclass of 7-dimensional complex filiform Leibniz algebras and their isomorphism classes <i>Sharifah Kartini Said Husain, Isamiddin S. Rakhimov</i>	1
New key exchange in Elliptic Curve based on Decomposition Problem <i>Hilyati Hanina Zazali, Wan Ainun Mior Othman</i>	7
Generators of Diagram Groups From Semigroup Presentation $P = \langle x, y, z \mid x = y, y = z, x = z \rangle$ Using Lifting Methods <i>Yusuf Gheisari and Abd Ghafur Bin Ahmad</i>	11
Finite Field Basis Conversion <i>Intan Muchtadi-Alamsyah, Marisa W. Paryasto, Muhammad Hafiz Khusyairi</i>	15
On Commutative Group of N-homomorphisms <i>Indah Emilia Wijayanti</i>	19
The f -chromatic indexes of wheel-like graphs <i>Adiwijaya^{1,2}, A.N.M. Salman¹, D. Suprijanto¹, E.T. Baskoro¹</i>	24
Structure of Incidence Algebras of Locally Finite Partially Ordered Set <i>Gustina Efriyanti¹ and Irawati²</i>	28
Large graphs diameter two from the smaller ones <i>Yus Mochamad Cholily</i>	37
On Module Classes Closed Under Submodules , Factor Modules and Direct Sums <i>Suwarno Ariswoyo and Elvina Herawaty</i>	40
Multinomial Option Pricing with Pseudoinverse Matrix <i>Abdurakhman</i>	44
Solving Traveling Salesman Problem Using A Hibryd of Evolution Strategies and Lin-Kernighan Algorithm <i>Nurmaulidar</i>	46
Analysis	54 - 84
Generation of Rainfall Sequence using Fourier Series <i>Norzaida Abas^{1*}, Zalina Mohd Daud¹, Fadhilah Yusof²</i>	55
On Extremal Properties for Certain Classes of Analytic Functions <i>Shaharuddin Cik Soh and Daud Mohamad</i>	61
The Area of the Region Enclosed by Bezier Curve <i>¹Normi binti Abdul Hadi*, ²Abdul Halmie bin Muhamad</i>	66
The Application of Regulated Function on the Multiplication of Two Henstock Integrable Functions <i>Christiana Rini Indrati</i>	70
Riesz potential and the generalized Morrey spaces with growth measures <i>Idha Sihwaningrum^{1*}, Hendra Gunawan², Yudi Soeharyadi³, Wono Setya Budhi¹</i>	75
An Estimation of Exponential Sums Associated With A Sextic Form <i>Sapar S.H¹, Mohd Atan K.A²</i>	79

Computational Mathematics	85 - 180
A Zero-dissipative Runge-Kutta-Nyström Method with Minimal Phase-lag for Oscillatory Problems <i>Norazak Senu^{1*}, Mohamed Suleiman², Fudziah Ismail³ and Mohamed Othman⁴</i>	85
New Multi-step Runge-Kutta method <i>O. Y. Ababneh^{1*}, R. Ahmad², E. S. Ismail³</i>	91
An efficient parallel implementation of Markov clustering algorithm for large-scale protein-protein interaction networks that uses MPI <i>Alhadi Bustamam^{1, 2, *}, Muhammad Shoaib Sehgal¹, Nicholas Hamilton¹, Simon Wong¹, Mark A Ragan¹, Kevin Burrage^{1, 3}</i>	94
Implementation Of Parallel Computational Tools For The Curing Simulation Of Thermoset Composites Using The Two Dimension Age Algorithm <i>Amna Abdurrahman¹, Ahmad Kamal bin Zulkifle², Norma Alias³, and Ishak Hashim⁴</i>	102
Optimizing of Text Retrieval: A similarity level by keyword competition in Genetic Algorithm (GA) <i>Poltak Sihombing¹</i>	109
The mutation and crossover effect in genetic algorithm to determine the similarity level of document retrieval <i>Poltak Sihombing¹</i>	117
Analysis and Evaluation RC4 Algorithm for Data Encryption <i>Ratna Aisuwarya¹, Rahmi Eka Putri²</i>	124
Efficient Differential Equation Solvers for Fluid Modeling in Interactive Surgical Drilling Simulations based on GPU Computations <i>Sugeng Rianto^{1*}, Ling Li²</i>	132
Developing Secant Method for Solving Nonlinear Equations <i>Taufiq Iskandar¹, Marlan², Sanggam P. Gultom³, Herman Mawengkang⁴</i>	137
Visualization of XLM-based Geochemical Data using SVG and ASP.NET <i>Nizamuddin^{1*} and Hidehiro Ishizuka²</i>	151
Application of Backtracking Algorithm On the Knight's Tour Game Using Pascal <i>Senja Omega Puspita, Budi Rudianto</i>	156
A Mixed Integer Linear Programming Model for Capacity Selection Problem in Logistics Networks Optimization <i>Muhammad Izman Herdiansyah¹</i>	164
Analysis and Implementation of the New Student Acceptance System On 2 nd Padang's Educational Council in On-Line Method By Using Applserver 2.5.7 Application <i>Mohammad Hafiz Hersyah</i>	175
Statistic	181 - 390
Goodness of Fit Test for the EEG Distribution by Using the Empirical Laplace Transform <i>Prasong Kitdamrongsuk^{1*}, Pachitjanut Siripanich²</i>	181
Conditional Maximum Likelihood Estimator for Incomplete Longitudinal Data <i>Juthaphorn Saekho^{1*}, Pachitjanut Siripanich²</i>	186
Autocorrelation correction in a simultaneous equations model <i>Kerativibooly W.^{1*}, Jitthavech J., Lorchirachoonkul V.</i>	192
Construction of Weights for the Estimation of Regression Coefficients with Outliers <i>Pimpan Ampanthong¹, Prachoom Suwattee²</i>	198
Goodness of Fit for the Poisson Distribution Based on Sample Skewness <i>Manad Khamkong[*] and Pachitjanut Siripanich</i>	204
Estimators of Errors of Kernel Density Estimator <i>Manachai Rodchuen^{1*}, Prachoom Suwattee¹</i>	209

Using Data Envelopment Analysis (DEA) Model to Measure The Relative Efficiency Of Higher Education Institutions <i>¹Zalina Zahid, ¹Rasimah Aripin and ²Mohd. Nasir Taib</i>	216
Detecting Students at Risk of Failing Using Decision Tree <i>Haliza Hasan[*], Rasimah Aripin & Sharifah Sakinah Syed Hassan Aidid</i>	222
Assessing The Performance of University Departments Using Data Envelopment Analysis <i>Nordin Hj Mohamad^{1*} and Fatimah Said²</i>	228
Mean-VaR Portfolio Optimization Under CAPM by Non Constant Volatility in Return Market <i>Sukono¹, Subanar² & Dedi Rosadi³</i>	238
Credit Risk Measurement for a Single Facility in Banking <i>Sukono</i>	242
Asymptotic Properties of a Generalized Renewal Reward Process <i>Suyono</i>	248
Distribution of a Generalized Renewal Reward Process <i>Suyono¹⁾ and Subanar²⁾</i>	252
Bayesian Updating Reservoir Simulation Models <i>Sutawanir Darwis^{1*}, Agus Yodi Gunawan², Sri Wahyuningsih¹, Nurtiti Sunusi¹, Aceng Komarudin Mutaqin¹</i>	256
Developing Bayesian inferential method for basic reproduction number in epidemic models under uncertainty <i>Dapot Situngkir, Abdul Latif Hasibuan, Daswati Sigalingging, Debora S.Parapat, Erwin Sidabalok, Herman Mawengkang</i>	261
Two stage mixed integer nonlinear stochastic programming model for solving a superstructure synthesis water networks optimization problem under uncertainty <i>Mujio, Evi Yanti Lubis, Harris H.Simamora, Herbin Manurung, Januasi Simarmata, Herman Mawengkang</i>	270
Application of Cluster Analysis In Classification of Tourist Destinations In Sabang <i>¹Asep Rusyana, ²Evi Ramadhani, ³Suhartono</i>	278
Simulation of Agent Based Model on Jakarta Stock Exchange (BEJ) <i>Afdal Mazni</i>	281
Modeling and Analyzing Dependent Categorical Data <i>Georgina M. Tinungki</i>	296
Life Insurance Product Valuation under Binomial Framework <i>Danang Teguh Qoyyimi[*], Danardono, Abdurakhman¹</i>	304
Time-dependent Strata as an alternative for Time-dependent Covariates in Survival Analysis <i>Danardono¹</i>	313
Statistical Analysis on Determination of Optimum Condition for Pra-esterification of Crude Palm Oil with Methanol to Biodiesel Using Sulfuric Acid Catalyst <i>Sawaluddin^{1*}, Tirenah Bahnur Siregar², Suwarno Ariswoyo³</i>	318
A Linear Mixed Model for Two-Dimensional Competition between Neighbouring Trees in Forestry Trials: Model Testing in a Plantation of Maritime Pine (<i>Pinus pinaster</i> Ait.) in Western Australia <i>Dadan Kusnandar^{1*}, N.W. Galwey²</i>	322
Reducing Fuzzy Relations of Fuzzy Time Series Model Using QR Factorization Method and Its Application to Forecasting Interest Rate of Bank Indonesia Certificate <i>Agus Maman Abadi¹, Subanar², Widodo³, Samsubar Saleh⁴</i>	328
An Adaptive Sensitivity-Based Linear Learning Method Algorithm for Data Classification <i>Zaenal Arifin¹⁾, M. Isa Irawan²⁾</i>	333
The Study of Sensitivity of Radial Basis Probabilistic Neural Network <i>Hasanuddin^{1*}, M. Isa Irawan²</i>	344

Recurrence Time Modeling for Earthquake Prediction <i>Surianto¹, Sutawanir Darwis², Aceng Komaruddin Mutaqin³</i>	350
Stochastic Project Scheduling <i>Johannes P. Sitanggang, Budi Irwansyah, Risna Helvida, Yuliani Nasution, Syafaruddin, Herman Mawengkang</i>	353
Mewma And Ewma Quality Control Charts, And Its Multiobjective Optimization <i>Arrival Rince Putri¹</i>	365
Characteristic of Poisson Distribution <i>Nova Noliza Bakar</i>	369
The Comparison of Forecasting of Jakarta Composite Index by Using Exponential Smoothing and Arima Method <i>Puspa Amelia¹⁾ and Maiyastri²⁾</i>	375
Representing Students Perspectives to Television Channels by Using Multidimensional Scaling Method (Case study in Department of Mathematics Andalas University) <i>Primawati¹*, Izzati Rahmi, HG²</i>	382
Latin Square Arrangement for Taste-Panel Experiments <i>Arisman Adnan</i>	387
Mathematics Education	391 - 445
Improving Student Academic Performance by An Application of Data Mining Techniques <i>Sajadin Sembiring¹*, Abdullah Embong², Mohd. Azwan Mohamad³, Muhammad Furqan⁴</i>	391
To consider the education of mathematics in different educational levels and it's more effective learning methods <i>Manochehr Kazemi¹ and Hassan Naraghi²</i>	396
How to Make the Teaching of Abstract Algebra Interesting? <i>Abdul Razak Salleh</i>	398
Matriculation Students' Metacognitive Awareness And Achievement In Mathematical Problem Solving <i>Effandi Zakaria¹*, Zainah Yazid²</i>	404
How The Indian Vedic Mathematics on Method of Calculation is still Relevant for Children in Today's Era of Mental Computation <i>Rita Desfitri¹ 2</i>	410
Prioritization of Factors to Further Studies in a University Using AHP <i>Yuzainee Bte Md Yusoff¹, Norngainy Bt Mohd Tawil²</i>	417
Contribution of Games to the Student Interest on Mathematics Lesson <i>Syukma Netti¹, Niniwati²</i>	423
Implementation of Creative Problem Solving Method by Using Media Computer to Improve Students' Achievement in Mathematics (Classroom Action Research at SMP 13 Pekanbaru) <i>¹Yenita Roza, ²Maida Deli</i>	425
Mathematic for Senior High School It's Problems and Challenges <i>Ali Asmar</i>	429
Errors in College Level Theorem Proving <i>I Made Arnawa</i>	432
Development of Model - Mathematics Learning Based of Interpersonal Intelligences For Student Class VII in Padang <i>Atus Amadi Putra</i>	437
Teaching Mathematics Through Cooperative Learning and Using ICT <i>Hendra Syarifuddin</i>	441

Improving Students' Activities And Mathematics Achievement Of SMPN 26 Padang Through Cooperative Learning; STAD Type <i>Mirna</i>	443
Applied Mathematics	446 - 960
The modification of steepest descent method On portfolio selection <i>Yosza Dasril¹, Goh Khang Wen², & Ismail Bin Mohd³</i>	446
Feedback Control and Magnetic Field Effects on Marangoni Instability in a Micropolar Fluid <i>Mohd Nasir Mahmud¹, Zainol Mustafa², and Ishak Hashim^{2*}</i>	452
A Mathematical Models of Lower Limb Using Kane's Method: An Application to Walking Movement While Carrying Load <i>Fazrolrozi & A. S. Rambely</i>	458
Mixed convection boundary layer flow towards a vertical plate embedded in a porous medium <i>Norfifah Bachok^{1*}, Anuar Ishak², Roslinda Nazar³, Ioan Pop⁴</i>	465
A New Number Representation for Faster Elliptic Curve Scalar Multiplication <i>Mohamad Rushdan Md Said[*], Abdul Wahed M. Ismail</i>	472
Effect of Non-Uniform Temperature Gradient and Magnetic Field on Marangoni Convection in a Micropolar Fluid <i>^{1,*}Melviana Johnson Fu, ²Norihan Md. Arifin, ³Mohd Noor Saad, ⁴Roslinda Mohd Nazar</i>	477
MHD stagnation-point flow towards a stretching sheet with induced magnetic field <i>Fadzilah Md Ali^{1*}, Roslinda Mohd Nazar², Norihan Md Arifin¹, Ioan Pop³</i>	483
Multiple Intelligence Profiling Analysis of People with Epilepsy for Job Placement Purposes <i>Siti Rahmah Awang^{1*}, Rasimah Aripin², Md. Hanip Rafia³</i>	489
Mathematical Modeling of the Upper Limb Movement using Kane's Equation <i>Sharifah Alwiah AbdulRahman^{1*}, Azmin Sham Rambely², Rokiah Rozita Ahmad³</i>	496
Pixel area validation of segmented malignant tumors in digital mammographic images <i>Rohana Embong^{1*}, Wan Eny Zarina W.Abd. Rahman¹, Tahir Ahmad², Rozi Mahmud³, Arsmah Ibrahim¹, Zainab Abu Bakar¹, Md Saion Salikin⁴</i>	500
Structural Similarity Measure for Mathematics Assessment Marking Engine <i>¹Arsmah Ibrahim, ²Zainab Abu Bakar, ³Nuru'l – 'Izzah Othman</i>	504
A Decision Making Model Based on Consistent Fuzzy Preference Relations and Generalized TOPSIS <i>Nor Hanimah Kamis, Daud Mohamad and Nor Hashimah Sulaiman</i>	510
The Study of Delivery Response Time: A Gap Analysis Approach <i>¹Mohd Sahar Sauian, ²Mohamad Yazid Shuaib</i>	517
Mathematical <i>Musyarakah</i> Model in Managing Islamic Investment Between Two Parties Using Two Profit Sharing Rates <i>Maheran Mohd Jaffar</i>	521
A Fuzzy Ruled Based Model for Stock Selection Advisor System <i>Daud Mohamad, Noorhar Jiana Haryanti Mohd Saad</i>	526
Performance of A Low-Cost Pcs in Edge Detection of Breast Tumor in Digital Mammomograms Using Wavelet Modulus Maxima <i>Arsmah Ibrahim¹, Norma Alias², Hanifah Sulaiman¹, Mohd Idris Jayes¹, Khairil Iskandar Othman¹, Md Saion Salikin³</i>	531
A Metaheuristics Approach for the Inventory Routing problem <i>Huda Zuhrah Ab Halim[*], Nur Arina Bazilah Aziz, Noor Hasnah Moin</i>	536
Single-vendor single-buyer model under linearly decreasing demand <i>Supadi, S.S^{1*}, Omar, M.²</i>	541

Catastrophe Reinsurance <i>Somayeh Nik Manesh^{1*}, Dr. Noor Azlinna Azizan²</i>	547
Determination of Work Done by a Female Student While Carrying Backpack with Different Loads <i>¹Nor Atikah Ab Ghani and ²Azmin Sham Rambely</i>	557
Linearity in Between House Price and Annual Management Fund in High-Rise Residential in Kuala Lumpur, Malaysia <i>N. M. Tawil¹, M. N. Daud³, A.I. Che-Ani², N.A. Goh², M.F.M. Zain²</i>	562
Fuzzy Conjoint Analysis of Influence Factors in High-Rise Residential Price And Management Fund In Kuala Lumpur, Malaysia <i>N. Mohd-Tawil¹, A.I. Che-Ani², Amiruddin Ismail³, M.M. Tahir², M. Jamil²</i>	569
Value-based Total Performance Measurement: A Preliminary Review <i>Z. Mustafa¹, N. R. M. Suradi¹, W. R. Ismail¹, K. A. M. Ali¹, Z. M. Ali¹, M. Abdullah¹, F. Idris², M. R. Ab Hamid³</i>	573
Simple Version of the Linear Business Cycle Model <i>Anton Abdulbasah Kamil^{1*} and Adam Baharum^{2*}</i>	577
Modelling The Relationships Between US and Selected Asian Stock Markets <i>¹Mohd Tahir Ismail*, ²Rosmanjawati Abdul Rahman</i>	581
Diffusive Logistic Equations with Single Time Delay and Variable Time Impulses <i>Jalina Widjaja</i>	587
Developing a Model for Estimating Emission Caused by Vehicles at a Junction of Main Street <i>Agus Salim Harahap, Herman Mawengkang</i>	590
The Pricing of European Type Asian Options with Geometric Averages <i>Mila Novita¹⁾, Erlina Natalia²⁾, Novita Dwi Hapsari³⁾, Widia Desrianti⁴⁾</i>	594
The Use of Genetic Algorithm for Selecting The Eigenvectors In Pca Method For Face Image Recognition <i>Suryadi MT, Yudi Satria, Helmiyati, Rahmi Rusin</i>	601
Model of Portfolio Insurance with Optimal Strike Price <i>Novriana Sumarti*, Muhammad Syamsuddin, Rieske Hadiani</i>	608
Cheapest Insertion - Convex Hull Approach to Euclidean TSP <i>Abdulah Fajar*, Nanna Suryana Herman², Nur Azman Abu³</i>	612
A Dynamic Nelson-Siegel Model for Indonesian Government Bond Yield Rates <i>R. Rahmawati*,^{1,2} M. Syamsuddin^{1,3}, R. Hadiani^{1,4}, S. Afriani^{1,5}, F. Damayanti^{1,6}, A.P. Wulandari^{1,7}</i>	617
Generalized reduced gradient method For earthquake resistant of Foundations <i>Abdul Hakam</i>	624
Modeling Dependence of Claim Amount between Different Claim Types using Copula <i>Yulia Resti¹⁾, Noriszura Ismail²⁾ & Saiful Hafizah Jaaman²⁾</i>	630
The Role of Mathematics to Determine Kiblat Direction <i>Akhsanul In'am</i>	634
Dialog of Features Characteristics by Using Venn Diagram for Object Detecti <i>Rahmadi Kurnia</i>	638
Comparison of Vector Stream Cipher Based on Modification of Chebyshev Polynomial Function <i>Santi Indarjani¹, Bety Hayat Susanti² and Jumiaty³</i>	645
Workforce planning problem with demand changes <i>Rustam Effendi Pasaribu, Rusly Siagian, Sabar, Saprida Montaria, Zulkifli, Herman Mawengkang</i>	650
On solving the Plant Cycle Location Problem <i>Henry Nainggolan, Misnawati, Nurtaito Sianturi, Rizky Ismalinda, Roslinawati, Herman Mawengkang</i>	663
A mixed integer linear programming model for multi-item inventory problems <i>Toha, Khairunisa Siregar, Lin Risnawati, Lord Byron Silalahi, Marsito, Herman Mawengkang,</i>	665

Integer programming model for operational aircraft maintenance routing problem with side constraints <i>Suhardi¹, Simon Petrus Sebayang², Benar³, Ardianta⁴, and Herman Mawengkang⁵</i>	673
Developing Mathematical Models for Analyzing The Performance Of A Growing Team <i>Yulidar¹, Zunaida Sitorus², Makmur Tarigan³, Satriawan Taruna⁴, Gim Tarigan⁵, and Herman Mawengkang⁵</i>	681
Discovering Relationship of Single Word Entities from Weblogs <i>Taufik Fuadi Abidin, Rasudin Abubakar, Alim Misbullah, Jufri Wahyudi</i>	695
Evaluation on Fitness Assignment Methods for Multi-objective Examination Timetabling Problems <i>Taufiq Abdul Gani^a, Nurmaulidar^b, Ahamad Tajudin Khader^c</i>	700
Face Recognition Using Smooth Support Vector Machine Based On Eigenfaces <i>^{1,2*}Muhammad Furqan, ²Abdullah Embong, ²Suryanti Awang, ²Santi W. Purnami, ²Sajadin Sembiring</i>	709
Mathematical Simulation Circuit for the Scale Model of Geometry Normalized Electromagnetic System (GNES) Equipment <i>*Muhammad Syukri Surbakit^{1,2}, ^aMohd Zubir Mat Jafri², ⁵Lim Hwee San², [^]Norhaslinda Mohamed-Tahrin²</i>	716
Handling Missing Values in Multiclass Multisurface Proximal Support Vector Machines <i>Taufik Edy Sutanto</i>	719
Traffic Flow Simulation on Simple Continuum Model <i>Sri Mardiyati¹ and Helen Burhan²</i>	725
Prototyping of a Gradient-Based Edge Detection Algorithm Design <i>Indra Yasri*</i>	729
A Mixed Integer Linear Programming Model for Capacity Selection Problem in Logistics Networks Optimisation <i>Muhammad Izman Herdiansyah¹</i>	733
Modeling Contagiousness of Diarrhea Diseases: A Spatial Probit Model <i>Yusep Suparman^a, Imam Munandar Fajari, Gatot Riwi Setyanto</i>	744
Model of Growth Population by Modification of Malthus and Gompertz Growth Models <i>Georgina M. Tinungki</i>	747
Stability Analysis and Maximum Profit of Wangersky-Cunningham Population Model with Time Delay and Constant Effort of Harvesting <i>Syamsuddin Toaha</i>	752
Fuzzy correlation to construct interaction function in group decision making <i>Marwan Harahap</i>	762
A Tabu Search with EST-SPT Algorithm for the Job Shop Scheduling Problem <i>Opim Salim Sitompul</i>	768
Finite Volume Method Based Analysis of Gas Flow in Two-Stroke Engine <i>Tulus</i>	773
An Improved Direct Feasible Search Approach for Solving Mixed-Integer Non Linear Programming Problems <i>Elly Rosmaini, Herman Mawengkang</i>	778
Optimization Methods for The Second Order Multiresponse Surface Model of Mixture Designs <i>¹Ruslan, ²Susanti L., ²Purhadi, ²Sony S</i>	782
The Application Of Classification Tree Method To Determine The Profile of Indonesian People Based on The Factors Which Significantly Influence The Attitudes Toward Avian Influenza <i>Rianti Setiadi</i>	786
Binary Structural Equation Model And Its Application To Find The Relationship Between "Knowledge About The Spread Of Avian Influenza" And "Preventive Actions Taken Toward Avian Influenza" <i>Rianti Setiadi</i>	792
Applications of Fuzzy Number Max-Plus Eigenvalues on Queuing Networks with Fuzzy Activity Times <i>M. Andy Rudhito^{1*}, Sri Wahyuni², Ari Suparwanto³, and F. Susilo⁴</i>	799

Developing a Direct Search Algorithm For Solving the Capacitated Open Vehicle Routing Problem <i>Tarno, Abdillah, Sudarman, Esmine Simatupang, Loide Naiborhu, Seprianti Harahap, Herman Mawengkang</i>	820
An Improved Feasible Neighbourhood Search Approach For Solving The Capacitated Facility Location Problem <i>Pramana, Tiopan Rahmat Siregar, Adil Pangaribuan, Sudarman Siringoringo, Indra Maryanti, Herman Mawengkang</i>	827
Finding Shortest Path in Networks with Uncertain Arc Length <i>Syamsul Qomar, Dewi Monalisa, Agus Budianto, Rosmartina, Lasma Nurhaida Silitonga, Yulis Purwaningsih, Herman Mawengkang</i>	833
Developing a Direct Search Algorithm For Solving the Capacitated Open Vehicle Routing Problem <i>Sudarman, Tarno, Abdillah, Esmine Simatupang, Loide Naiborhu, Seprianti Harahap, Herman Mawengkang</i>	838
Development of Solow Growth Model <i>Hamidah Nasution¹ and Herman Mawengkang^{2,1} Matematika MIPA UNIMED, ²</i>	845
Deciding Hub Location in a Communication Network <i>Ummi Habibah, Rosimanidar, Purnawanto, Herman Mawengkang</i>	849
A Constrained Based Approach for Handling The Multi-Period Single-Sourcing Problem <i>Afnaria, Herman Mawengkang</i>	860
A Goal Programming Approach for a Class of Possibilistic Portfolio Selection Model <i>Rina Filia Sari, and Herman Mawengkang</i>	874
Optimization Model for Land Management Problems under Uncertainty <i>Almira Amir</i>	883
An Optimization Model for Irrigation Water Distribution Networks <i>Alfred Hasiholan Silalahi, Rahmanan Dalimunthe, Isabella Bangun, Tiramah Simanjuntak, Surya Ningsih, Herman Mawengkang</i>	889
Developing a constrained search approach for solving systems of nonlinear equations <i>Pasukat Sembiring</i>	894
On The Sufficient Condition for Solvability of Infinite Horizon LQ Problem Subject to DAE Systems <i>Muhafzan¹</i>	903
Experimental Modelling of Domestic Tourist In West Sumatera by Using Separate Spline Function <i>Susila Bahri</i>	908
Structural Equation Approach for Non-Normality Data: With Reference to Modeling of Health Index <i>Ferra Yanuar^{1*}, Kamarulzaman Ibrahim², Abdul Aziz Jemain³</i>	910
Size Ramsey Number for P_3 and T_3 <i>Des Welyyanti</i>	917
Arranging The Examination Schedule By Using Graph Coloring Algorithm <i>¹Budi Rahmadya, ²Narwen</i>	919
Monitoring Mangrove Rehabilitation In Tsunami Affected Area Using High Resolution Satellite Images <i>Muzailin Affan</i>	927
Numerical Solution of Iron Corrosion Problem Based on Condensation Chemical Property <i>¹Arid Fatahillah, ²Basuki Widodo</i>	930
Constructing Model for Survival Data by the Makeham-like Unproportional Hazard <i>Adhitya Ronnie Effendie^{1*}, Danardono², Subanar³</i>	935
Comparative Analysis of Fuzzy Logic and Multiple Linear Regression On Forecasting of Brand Switching Based on the Value of Consumer Dissatisfaction, the Characteristics of Product Categories And Explore the Needs of Variation <i>Resti Athayani, Budi Rudianto</i>	940

The Implementation Of Finite Difference Method To Simulation Of Transient Temperature Distribution Of Three-Dimensional Heat Conduction In Cartesian Coordinate System <i>Hufri</i>	949
Power Series Solution Of Non-Linear First Order Differential Equation Systems <i>Zulakmal</i>	955
Model Reduction Using LMIs <i>Jenizon</i>	958
Identification Algorithm For Temporal Logics <i>Yahma Wisnani</i>	960
An edge consecutive edge magic total labeling on some classes of tree <i>Kiki A. Sugeng and Denny R. Silaban</i>	966
On (a,d)-vertex antimagic labeling of circulant graphs <i>Bong N. Herawati and Kiki A. Sugeng</i>	970
Determination of The Sex of Hawksbill Sea Turtle (<i>Eretmochelys imbricata</i>) by Using Logistic Regression <i>Hazmira Yozza, Hilda Yohana, Izzati Rahmi HG, Kurniadi Ilham</i>	974

Determination of The Sex of Hawksbill Sea Turtle (*Eretmochelys imbricata*) by Using Logistic Regression

Hazmira Yozza, Hilda Yohana, Izzati Rahmi HG, Kurniadi Ilham

Abstract

Logistic regression is a statistical analysis used to predict probability of an occurrence by analyzing the relationship between explanatory variables and binary response variable. In this paper, logistic regression is used to determine important variables that enable us to predict the sex of a hawksbill sea turtle (*Eretmochelys imbricata*). It is found that there are two important variables, which are straight carapace width and vent to tail tip distance. The wider the carapace width the bigger the probability of a turtle for being a male. The shorter the distance between vent to tail tip, the bigger the probability.

Keyword : Logistic regression, binary response, Hawksbill sea turtle

1. Introduction

Hawksbill sea turtle is one of marine species that are considered as endangered species by The International Union for the Conservation of Nature and Natural Resources. The *hawksbill sea turtle* global population declines 80 percent or more during the past century and continuous decline is projected. The decline is primarily due to human exploitation for their meat, eggs and shell. Other threats include loss or degradation of nesting habitat from coastal development and beach armoring; disorientation of hatchlings by beachfront lighting; excessive nest predation by native and non-native predators; marine pollution, watercraft strikes, and incidental take from commercial fishing operations.

Efforts have been done to save the turtle from extinction, either naturally or artificially. Naturally, it is done by nest protection and eggs relocation. Artificially, it is done to increase the number of female turtle. One of this approach is done by selective rearing based on sex differentiation. By using this approach, the sex of the turtle can be detected earlier.

The problem faced is that turtle doesn't have external genital. To determine the sex, one have to kill the turtle, which is not allowed due to the endengared status of the turtle. Alternatively, we can develop a model based on morfometrics measurement of the turtle to determine the sex, which is a proven approach for determining the sex of other species, such as bird or fish. The aim of this research is to develop a model to determine the sex of the turtle based on morfometrics of the turtle.

2. Logistic Regression

Regression analysis is a statistical technique that serves as a basis of drawing inference about relationship among quantities, i.e a response variable and a set of explanatory variables. When the response variable is continuous, this relationship is often of the form:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon_i$$

which expresses that the fact that the response for the i th observation y_i , $i = 1, 2, \dots, n$ depends linearly on the values of k explanatory variables labeled x_1, x_2, \dots, x_k through unknown parameters $\beta_0, \beta_1, \dots, \beta_k$. The explanatory variables are assumed to be fixed and measured without error. The term ε_i is an observable random variable which represent the residual variation and assumed normally distributed with zero mean and

homogeneous variance, σ^2 .

In many practical situation, the response is basically binary, i.e. there are two possible outcomes that may be assigned as 0 (failure) or 1 (success). For binary data, the response from the i th observation, $i = 1, 2, \dots, n$, is a

proportion y_i/n_i , which is conveniently denoted by \tilde{p}_i , the probability of success. The model-building is done to determined the role of a set of explanatory variables x_1, x_2, \dots, x_p on the binary response. In addition, it may be a need to predict or, rather, estimate the probability of one of two spesific response outcome at a certain combination x_1, x_2, \dots, x_p .

One approach to modelling binary data, sadly encouraged by widespread availability of statistical software

for linear regression analysis, is to adopt the model $p_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon_i$ and

apply the method of least square to obtain those values $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$ for which :

$$\sum_i \left(\frac{y_i}{n_i} - p_i \right)^2 = \sum_i (\tilde{p}_i - \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki})^2$$

is minimised.

The use of ordinary least square in this binary situation will bring violation on two important assumptions, namely :

1. Distribution of the ε is discrete, and thus not normal

2. Error variance is not homogeneous

Another problem is of fundamental and concerns the fitted values, \hat{p}_i , under the model. The \hat{p}_i are totally unconstrained and can take any value, positive or negative, large or small, so that any linear combination of them can in principle lie anywhere in the range $(-\infty, \infty)$. Since fitted probabilities are obtained from $\tilde{p}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \dots + \hat{\beta}_k x_{ki}$, there can be no guarantee that the fitted values will lie in the interval (0,1).

Instead of using a linear model for the dependence of success probability on explanatory variables, the probability scale is first transformed from the range (0, 1) to $(-\infty, \infty)$. A linear model is then adopted for the transformed value of the success probability, a procedure which ensures that the fitted probabilities will lie between zero and one. A very popular and useful transformation is logistic transformation. The logistic transformation of a success probability p is $\log \{p/(1-p)\}$ which is written as $\text{logit}(p)$. Notice that $p/(1-p)$ is the odds of a success and so the logistic transformation of p is the log odds of the success. It is easily seen that any value of p in the range (0,1) corresponds to a value of $\text{logit}(p)$ in $(-\infty, \infty)$. The linear model based on this transformation is :

$$\text{logit}(p) = \log \left[\frac{p_i}{1-p_i} \right] = \beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki}$$

By using some rearrangement, it will be found that :

$$p_i = \frac{1}{1 + e^{-\eta_i}} \quad i=1,2,\dots,k$$

Here, the quantity

$$\eta_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki}$$

This model relates the probability of occurrence p_i against explanatory variables, which is sometimes called as risk factors. β_0 is called the "intercept" and $\beta_1, \beta_2, \beta_3$, and so on, are called the "regression coefficients" of x_1, x_2, x_3, \dots respectively. The intercept is the value of η when the value of all risk factors is zero (i.e., the value of η in an object with no risk factors). Each of the regression coefficients describes the size of the contribution of that risk factor. A positive regression coefficient means that the risk factor increases the probability of the outcome, while a negative regression coefficient means that risk factor decreases the probability of that outcome; a large regression coefficient means that the risk factor strongly influences the probability of that outcome; while a near-zero regression coefficient means that the risk factor has little influence on the probability of that outcome.

In order to fit a linear model to a given set of data, the $k+1$ parameters, $\beta_0, \beta_1, \beta_2, \dots, \beta_k$, have first to be estimated. These parameters are readily estimated using the methods of maximum likelihood. The likelihood function is given by :

$$L(\beta) = \prod_{i=1}^n \binom{n_i}{y_i} p_i^{y_i} (1-p_i)^{n_i-y_i} = \prod_{i=1}^n \binom{n_i}{y_i} \left(\frac{p_i}{1-p_i} \right)^{y_i} (1-p_i)^{n_i}$$

Now, the problem is to obtain those values $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$ which maximize $L(\beta)$ or equivalently $\log L(\beta)$

The logarithm of the likelihood function is :

$$\begin{aligned} \log_e L(\beta) &= \sum_{i=1}^n \log_e \left(\binom{n_i}{y_i} \right) \left(\frac{p_i}{1-p_i} \right)^{y_i} (1-p_i)^{n_i} \\ &= \sum_{i=1}^n \left\{ \log_e \left(\binom{n_i}{y_i} \right) + y_i \log_e \left(\frac{p_i}{1-p_i} \right) + n_i \log_e (1-p_i) \right\} \\ &= \sum_{i=1}^n \left\{ \log_e \left(\binom{n_i}{y_i} \right) + y_i \eta_i + n_i \log_e \left(1 - \frac{1}{1+e^{-\eta_i}} \right) \right\} \\ &= \sum_{i=1}^n \left\{ \log_e \left(\binom{n_i}{y_i} \right) + y_i \eta_i - n_i \log_e (1+e^{\eta_i}) \right\} \end{aligned}$$

The derivation of this log likelihood function with respect to the $k+1$ unknown β parameters are :

$$\frac{\partial \log_e L(\beta)}{\partial \beta_j} = \sum_{i=1}^n y_i x_{ji} - \sum_{i=1}^n n_i x_{ji} e^{\eta_i} (1+e^{\eta_i})^{-1} \quad j = 0, 1, 2, \dots, k$$

Evaluating these derivatives at $\hat{\beta}$ and equating them to zero gives a set of $k + 1$ non-linear equations in the unknown parameters $\hat{\beta}_j$. These can only be solved numerically. Another approach is using Fisher's method of scoring. This procedure is equivalent to iteratively weighted least square procedure.

Goodness of Fit of a Linear Logistic Regression

After fitting a model to a set of data, it is natural to enquire about the extent to which the fitted values of the response variable under the model compared with the observed values. If the agreement between the observation and the corresponding fitted values is good, the model may be acceptable. If not, the current form of the model will certainly not be acceptable and the model will need to be revised. This aspect of adequacy of a model is widely referred to as goodness of fit.

There are various statistics that have been proposed for assessing the goodness of fit of logistic regression model, analogous to those that are used in linear regression. Some of those statistics are deviance, R^2 , Pearson Chi-Square and CCR (Correct Classification Rate).

The statistic deviance measures the extent to which the current model deviates from the full model. The deviance in logistic regression corresponds SSE in linear regression. The deviance, D , is defined as :

$$D = -2 \log \frac{L_c}{L_F}$$

where L_c is the maximum likelihood under the current model and L_F is the maximum likelihood under the full model. Large values of D are encountered when L_c is small relative to L_F , indicating that the current model is a poor one. On the other hand, small values of D are obtained when L_c is similar to L_F , indicating that the current model is a good one.

Pearson Chi-Square statistics is defined by :

$$\chi^2 = \sum_{i=1}^n \frac{(y_i - n_i \hat{p}_i)^2}{n_i \hat{p}_i (1 - \hat{p}_i)}$$

Various forms of R^2 have been proposed for the logistic model. Magee (1990) proposed using :

$$R^2 = 1 - \{L_0 / L_F\}^{2/n}$$

If the objective is to predict whether a subject will or will not have the attribute of interest, a more meaningful measure of the goodness of the model would be the percentage of subject in the data set that are classified correctly. This measure usually referred as CCR (Correct Classification Error)

Hypothesis Testing in Logistic Regression

The significance of logistic model is tested by likelihood ratio test. The hypothesis is given by :

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1 : \beta_j \neq 0 ; j = 1, 2, \dots, k$$

Deviance is used as a test statistic of this test. It is defined as :

$$D = -2 \log \frac{L_0}{L_F}$$

where L_0 is the maximum likelihood under the null model (i.e. model without explanatory variables) and L_F is the maximum likelihood under the full model. This statistic has χ^2 distribution with k degrees of freedom.

In linear regression, t-test is used to assess the significance of each individual explanatory variables, when other explanatory variables are in the model. In logistic regression, this test is conducted by using Wald test. The hypothesis is given by :

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0$$

The test statistic of this test is called a Wald statistic and defined as :

$$W = \left(\frac{\hat{\beta}_j}{s_{\hat{\beta}_j}} \right)^2$$

Under H_0 , Wald statistic have χ^2 distribution with 1 degree of freedom.

Coefficient Interpretation

The interpretation of the β_j parameter estimates is as the additive effect on the log odds ratio for a unit change in the j th explanatory variable. In the case of a dichotomous explanatory variable, for instance gender, e^{β} is the estimate of the odds ratio of having the outcome for, say, males compared with females.

Binary Response Prediction

A common goal in linear logistic modelling is to predict the value of a binary response variable. The predicted response probability can be obtained directly by adding the value of the explanatory variables for a

new observation. This predicted probability may subsequently form the basis for allocating an object according to one of two groups.

To assign an individual to one of two groups on the basis of their predicted response probability, a threshold value π_0 has to be identified. An object is assigned to group 1 if their predicted response probability $\hat{p} < \pi_0$ and to group 2 if $\hat{p} \geq \pi_0$.

DATA AND METHODS

Data

In this research, we use the sex of the turtle as response variable. We assign 0 for female and 1 for male turtle. Set of explanatory variables consists of 17 variables, i.e :

X_1 = SCL (Straigh Carapace Length) X_{10} = CPW (Curve Plastron Width)
 X_2 = SCW (Straigh Carapace Width) X_{11} = BD (Body Depth)
 X_3 = CCL (Curve Carapace Length) X_{12} = BW (Body Weight)
 X_4 = CCW (Curve Carapace Width) X_{13} = TL (Tail Length)
 X_5 = HL (Head Length) X_{14} = VTDD (Vent to Tail Tip Distance)
 X_6 = HW (Head Width) X_{15} = VTBD (Vent to Tail Base Distance)
 X_7 = SPL (Straigh Plastron Length) X_{16} = VL (Vent Length)
 X_8 = SPW (Straigh Plastron Width) X_{17} = VW (Vent Width)
 X_9 = CPL (Curve Plastron Length)

Methods

The data used in this research is obtained from the morphometric measurement of 28 hatchling hawksbill sea turtles. The sex of the hatchling is determined after incision. It is found 8 male turtles and 20 female turtles. Modelling the relationship between the sex of the turtle and their morfometrics is done by using logistic regression analysis. The hypotesis testing is done to see whether these 17 variables significantly influenced the sex of the turtle. Wald test is employed to determined the worth of the individual variables. The goodness of fit of the model is measured by Correct Classification Rate (CCR). The threshold value used to allocate turtle on the basis of its sex is 0.5.

RESULT AND DISCUSSION

The linear model for this relationship is :

$$\begin{aligned} \text{logit}(\hat{p}) = & -495.658 - 8.036 X_1 + 71.987 X_2 + 22.106 X_3 + 3.226 X_4 - 9.510 X_5 + 7.382 X_6 \\ & + 10.210 X_7 - 54.046 X_8 - 12.696 X_9 - 15.475 X_{10} + 13.528 X_{11} - 7.146 X_{12} \\ & - 0.446 X_{13} - 4.558 X_{14} + 0.386 X_{15} + 8.141 X_{16} - 3.731 X_{17} \end{aligned}$$

Hypothesis testing is done to see whether these 17 variables influence the probability of the turtle for being a male. The hypothesis is given by :

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{17} = 0$$

$$H_1 : \text{at least one } \beta_j \neq 0 ; j = 1, 2, \dots, 17$$

In this test, G-statistics is used as test statistic. This statistics follows χ^2 distribution with 17 degrees of freedom. From Chi-Square table, it is found that, for $\alpha=0.05$, $\chi^2_{0.05,17} = 27.6$. From the data it is obtained the G-statistic = 33.503, which is significant at a level 0.05. It means that H_0 is rejected in favor of H_1 , with the conclusion that at least of these 17 variables is statistically significant influenced the probability of turtle for being a male.

The worth of the individual explanatory variables will be determined by using Wald statistic. Following table shows the wald statistics of all variables.

Table 1. Wald Statistics of 17 Variables

Variable	B	S.E.	Wald	Df	p-value
SCL (X_1)	-8.036	96247.666	.000	1	1.000
SCW (X_2)	71.987	283451.074	.000	1	1.000
CCL (X_3)	22.106	282009.677	.000	1	1.000
CCW (X_4)	3.226	14365.357	.000	1	1.000
HL (X_5)	-9.510	710934.196	.000	1	1.000
HW (X_6)	7.382	1156375.140	.000	1	1.000
SPL (X_7)	10.210	75738.294	.000	1	1.000
SPW (X_8)	-54.046	1286099.118	.000	1	1.000

CPL (X_9)	-12.696	262134.799	.000	1	1.000
CPW(X_{10})	-15.475	59811.425	.000	1	1.000
BD (X_{11})	13.528	194124.844	.000	1	1.000
BW (X_{12})	-7.146	402535.434	.000	1	1.000
TL (X_{13})	-.446	6498.476	.000	1	1.000
VTTD (X_{14})	-4.558	27930.600	.000	1	1.000
VTBD (X_{16})	.386	3117.184	.000	1	1.000
VL (X_{17})	8.141	101137.362	.000	1	1.000
VW (X_{18})	-3.731	47161.244	.000	1	1.000
Constant	-495.658	17309654.255	.000	1	1.000

From the above table, it can be seen that no variables contributes significantly in the logistic model. This result is different from the result obtained before. It might be caused by the highly correlation among variables.

The best model is selected by using forward selection method. The following table shows the result of forward selection for this data.

Table 2. Forward Selection Result

	B	S.E.	Wald	Df	Sig.
1 VTTD (X_{14})	-.174	.079	4.934	1	.026
Constant	18.343	8.563	4.589	1	.032
2 SCW (X_2)	2.866	1.546	3.434	1	.044
VTTD (X_{14})	-.266	.107	6.150	1	.013
Constant	-31.451	26.279	1.432	1	.231

The best linear logistic model is :

$$\text{logit}(\hat{p}) = -31.451 + 2.866 X_2 - 0.266 X_{14}$$

From this model, we find two variables that contribute to distinguish the sex of the turtle, namely straight carapace width (X_2) and Vent-to-tail-tip distance (X_{14}). By measuring these variables, the grouping of turtle based on its sex can be done easily and efficiently.

Futhermore, the significance of this final model is tested. The G-statistic of this test is 14.008, which is greater than $\chi^2_{0.05,2} = 5.99$. Therefore it can be concluded that this model is significance at a level 0.05. Wald test that is conducted to determine the significance of each variables also gives the same result.

This best model can be written in different form :

$$\hat{p} = \frac{1}{1 + e^{-(31.451 + 2.866 X_2 - 0.266 X_{14})}} = \frac{1}{1 + e^{(31.451 - 2.866 X_2 + 0.266 X_{14})}}$$

This model can be used as a basis to predict sex of the turtle. The turtle is predicted as a male if $\hat{p} \geq 0.5$ and as a female if $\hat{p} < 0.5$. Here is the result.

Table 3. Turtle Classification

Observation		Predicted	
		Y	
Y	.00 1.00	.00	1.00
		18	2
		3	5

It can be seen from this table that 23 of 28 turtles is correctly classified. Thus, we can calculate the CCR = $23/28 \times 100\% = 82\%$

To interpret the significance of X_2 (SCW) in distiguish the sex of the turtle, odds ratio is used. Table 3 shows the odds ratio of SCW and VTTD in best model.

Table 4. The Odds Ratio in The Best Model

Variabel	$Exp(\beta)$
SCW (X_2)	17.559
VTTD (X_{14})	0.766

Table 3 shows that the odds ratio for SCW is 17.559. It means that for every 1 mm increment of the width of carapace, the increment of the risk (odds) of the observed turtle being a male is 17.559. In other words, the odds of turtle with certain carapace width is 17.559 time greater than turtles with carapace width 1 mm less.

The odds ratio for VTDD is 0.766. It means that for every 1 mm increment of the distance between vent to tail tip, the increment of the risk of the observed turtle being a male is 0.776. Or, In other word, the odds of turtle with certain value of vent-to tail-tip distance is 0.766 times greater than turtle with vent to tail tip distance 1 mm less. In other word, The shorter the distance between vent to tail tip, the bigger the probability.

CONCLUSION

Logistic modelling done to 28 young turtle found that there are two important variables, which are significantly influenced the probability of the turtle for being a male. These variables are straight carapace width and vent to tail tip distance. The wider the carapace width the bigger the probability of a turtle for being a male. The shorter the distance between vent to tail tip, the bigger the probability.

REFERENCE

- Agresti, A. 1990. *Categorical Data Analysis*. John Wiley & Sons, Inc., Canada
- Collet, D. 1998. *Modelling Binary Data*. Chapman & Hall, Canada
- Gaspersz, V. 1995. *Teknik Analisis dalam Penelitian Percobaan Jilid 2*. Penerbit Tarsito, Bandung
- Grahame, M. 2006 *Hawksbill Sea Turtle (Eretmochelys imbricata)* Massachusetts
<http://www.cdu.edu.au/ser/GrahameWebbProfile.htm>. May 8th 2009
- Kleinbaum, DG. 1994. *Logistic Regression : A Self Learning Text*. Springer, New York
- McCullagh, P and JA Nelder. 1989. *Generalized Linear Models 2nd edition*. Chapman and Hall, London
- Myers, RH. 1990. *Classical and Modern Regression with Application 2nd edition*. PWS Kent Publishing Company, Boston
- Ryan, TP. 1997. *Modern Regression Methods*. John Wiley & Sons, Inc., New York
- Yozza, H and I. Rahmi. 2006. *Penetapan Nilai Ambang (Threshold) Pemotongan P_i untuk Menduga Respons Biner pada Analisis Regresi Logistik*. HEDS Project