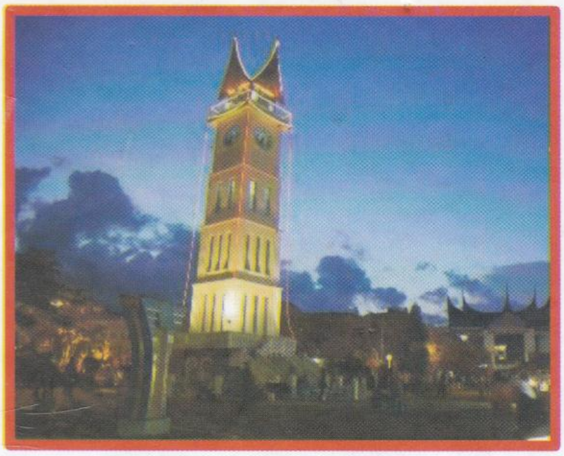
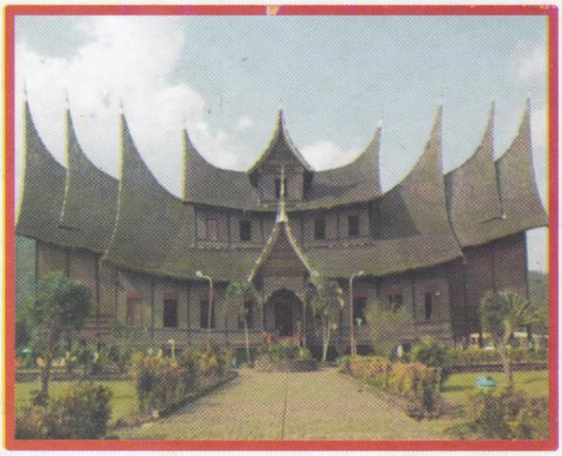
PROCEEDING

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

ICMSA 2009

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





Organized by :

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

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

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

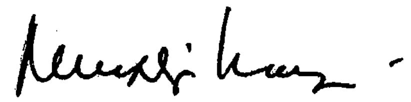
Message from Rector Andalas University

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/Accomodation	:	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

Computational Mathematics 85 - 180

Statistic 181 - 390

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 Resourses. 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 aprroach 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:

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

which expresses that the fact that the response for the *i*th observation yi, i = 1,2,...,n depens linearly on the values of k explanatory variables labbeled x_1, x_2, ..., x_k through unknown parameters β_0, β_1, ..., β_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_i^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,...,n, is a proportion y_i/n_i , which is conveniently denoted by $\widetilde{p}_i$, the probability of success. The model-building is done to determined the role of a set of explanatory variables x1, x2, ..., xp 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 x1, x2, ..., xp.

One approach to modelling binary data, sadly encouraged by widespread availability of statistical software for linear regression analysis, is to adopt the model $p_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \varepsilon_i$ and apply the method of least square to obtain those values $\hat{\beta}_0, \hat{\beta}_1, \cdots, \hat{\beta}_k$ for which :

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

is minimised.

The use of ordinary least square in this binary situation will bring violation on two important assumpsions, 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 $\hat{\beta}_j$ 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 probablities are obtained from $\tilde{p}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \cdots + \hat{\beta}_k x_{ki}$, 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 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. In is easily seen that any value of p in the range (0,1) corresponds to a value of logit (p) in $(-\infty,\infty)$. The linear model based on this transformation is :

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

By using some rearrangement, it will be found that :

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

Here, the quantity

$$\eta_i = \beta_o + \beta_1 x_{1i} + \cdots + \beta_k x_{ki}$$

This model relates the probability of occurance p_i against explanatory variables, which is sometimes called as risk factors. β_0 is called the "intercept" and β_1, β_2, β_3, and so on, are called the "regression coefficients" of x_1, x_2, x_3,... 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, $\beta_0, \beta_1, \beta_2,\ldots,\beta_k$, 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 $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2,\ldots,\hat{\beta}_k$ which maximize L(β) or equivalently log L(β)

The logarithm of the likelihood function is :

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

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

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

Evaluating these derivatives at $\hat{\beta}$ and equating them to zero gives a set of k + 1 non-linear equations in the unknow parameters $\hat{\beta}_j$. These can only be solved numerically. Another approach is using Fisher's method of scoring. This procedure is equivalent to iteralively 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 corresponsding fitted values is good, the model may be acceptable. If not, the current form of the model will certanly 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 subjec 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)

Hyphotesis Testing in Logistic Regression

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

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

$H1 : \beta_j \neq 0$; j = 1, 2,..., 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 :

$H0 : \beta j = 0$

$H1 : \beta j \neq 0$

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 jth 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 dor 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 treshold 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_2 = SCW (*Straigh Carapace Width*)
X_3 = CCL (*Curve Carapace Length*)
X_4 = CCW (*Curve Carapace Width*)
X_5 = HL (*Head Length*)
X_6 = HW (*Head Width*)
X_7 = SPL (*Straigh Plastron Length*)
X_8 = SPW (*Straigh Plastron Width*)
X_9 = CPL (*Curve Plastron Length*)
X_{10} = CPW (*Curve Plastron Width*)
X_{11} = BD (*Body Depth*)
X_{12} = BW(*Body Weight*)
X_{13} = TL (*Tail Length*)
X_{14} = VTTD (*Vent to Tail Tip Distance*)
X_{15} = VTBD (*Vent to Tail Base Distance*)
X_{16} = VL (*Vent Length*)
X_{17} = VW (*Vent Width*)

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 treshold value used to allocate turtle on the basis of its sex is 0.5.

RESULT AND DISCUSSION

The linear model for this relatonship is :

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

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 :
$H_0 : \beta_1 = \beta_2 = \cdots = \beta_{17} = 0$
$H_1 :$ at least one $\beta_j \neq 0$; $j = 1,2,\ldots,17$

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 α=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 :

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

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 famale if $\hat{p} < 0.5$. Here is the result.

Table 3. Turtle Classification

Observation		Predicted	
		Y	
		.00	1.00
Y	.00	18	2
	1.00	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 x 100% = 82%

To interpret the significance of X2 (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(β)*
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 VTTD 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 turte 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 (Treshold) Pemotongan p_i untuk Menduga Respons Biner pada Analisis Regresi Logistik*. HEDS Project