M – 05

Longitudinal Tobit Regression

Modelling Stroke Patients With Trauma/Injury HeadTrauma

Evy Annisa Kartika S[1], Ismaini Zain[2], Vita Ratnasari[3]
[1] Dept Of Statistics, Institute Teknology Sepuluh Nopember
[2] Dept Of Statistics, Institute Teknology Sepuluh Nopember
[3] Dept Of Statistics, Institute Teknology Sepuluh Nopember
evy.annisa14@mhs.statistika.its.ac.id

***Abstract*—** Stroke is a major health problem for modern society. In the case of patients with stroke, a stroke patients often experience trauma / head injury due to falls in the bathroom, fell on the stairs or after waking. One way to determine the severity of head injuries in stroke patients is used of medical records Glasgow Coma Scale (GCS). But the scoring system at GCS alone sometimes has problems, such as pain medications under investigation and response variables measured on the visual analog, some patients will be reported "no pain" after a certain time. They could not say a lesser degree or higher from a sense of "no pain". To solve it, scoring at GCS will be given sensor at a certain value. Medical records on the severity of the trauma / injury done repeatedly over time. Time is very important and in the measurement of an event expected to be recorded properly. One approach that can be performed on longitudinal data which has a sensor point is to approach tobit longitudinal regression method. The response variable is the score on the Glasgow Coma Scale. The predictor variables having a significant effect is blood pressure (X_1), body temperature (X_2) and pulse (X_3).

Keywords: *Stroke, Longitudinal Tobit, Glasgow Coma Scale*

I. INTRODUCTION

Stroke is a major health problem for our modern society. Today, sroke increasingly becoming a serious malash facing almost all over the world. That is because the sudden stroke can result in death, physical disability and mental well childbearing age and the elderly [1]

In the case of patients with stroke, a stroke patients often experience trauma / head injury due to falls in the bathroom, fell on the stairs or after waking. Trauma / head injury in stroke patients require ongoing medical record to ensure further medical treatment. With minimal facilities, the scoring system becomes very necessary to assess the state of consciousness and prognosis of the patients experienced injuries. One way to determine the severity of head injuries in stroke patients is the use of medical records Glasgow Coma Scale (GCS). GCS itself based on the patient's response to eye opening, verbal function and various motoric functions of the various stimulus [2]. GCS scoring system on his own sometimes have problems, such as pain medications under investigation and response variables measured on the visual analog, some patients will be reported "no pain" after a certain time. They can not tell the lower levels of "no pain". To overcome this problem then scoring on the Glasgow Coma Scale will be given sensor on one particular value.

In some studies, the response variables of the specific research has the highest limit from time to time. The actual level of the highest scoring test takers can not be measured properly and this phenomenon is called the effect of the highest limit. Effect occurs when the upper limit in a study, test scale relative ease so that everyone receives a proportion of the maximum or near the maximum value can not be determined [3]. To handle data censored at the upper limit tobit developed and applied approach.

The approach taken is to use longitudinal regression tobit where medical records GCS were observed repeatedly during the 14 days when the patient is treated, will be modeled factors that affect the severity of head trauma patients. It is hoped that by knowing the factors that influence can be seen the severity of injury / trauma and can be given treatment in the healing of patients to the fullest

II. METHOD

A. Tobit Regression

Model tobit (Truncated Regression) or so-called censored regression is a regression analysis that is used to describe the relationship between the dependent variable and the independent variable where the

dependent variable scale mix. Tobit The term was first proposed by James Tobin in 1958 were used to analyze the expenditures of the households in the United States to buy the cars [4]. Modeling tobit formed by first assuming the existence of a relationship linear between Y^* with variabel X expressed by :

$$y_i^* = x_i^T \beta + \varepsilon_i \tag{1}$$

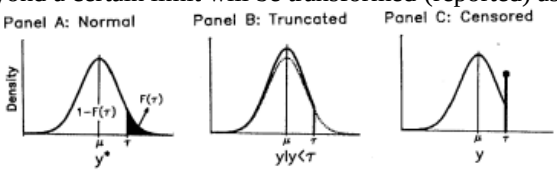
With y^* contains a latent variable that represents the value censored. When y^* with ceiling effects $y_i^* \geq \tau$, so the variabel respon (Y) cencored was definition likes:

$$y_i = \begin{cases} y_i^* = x_i^T \beta + \varepsilon_i & jika\ \ y_i < \tau \\ \tau & jika\ \ y_i \geq \tau \end{cases} \tag{2}$$

Where τ is threshold where $i = 1, 2, ..., n$ and n is the number of observations.

B. The Concept of Data Cencored

Sensor occurs when the observations were made on variables for the entire sample, but there are some observations on the response variable that has limited information. When a censored response variable, the value of observation beyond a certain limit will be transformed (reported) as a single value [5]



Long [8] explain that if you use all the data to the regression model to the data censored, will generate parameter values over estimate the intercept. Whereas if eliminating the observation that value is not known, will produce the coefficient parameters underestimate the slope and the intercept overestimate. And truncated cause correlations between predictor variables with residual, resulting in inconsistent estimates.

C. Longitudinal Data Analysis

Longitudinal data analysis is a study that refers to the observation that the measurement is repeated from time to time [6]. Such studies have appeared in various fields. In the field of medical study was conducted to measure the rehabilitation of stroke patients were observed at various times and circumstances. This study is very effective for studying the effects of time on certain characteristics.The main objective of the analysis was to determine the response of the middle value, expressed as a function of time. In general, longitudinal data set is written as follows:

$$\left(y_{ij}, t_{ij}\right) \quad i = 1, 2, ..., m; \ \ j = 1, 2, ..., n_i \tag{3}$$

Where:

y_{ij} : measurement to j (from n_i measurement) to object i (from m object)

t_{ij} : time measurement y_{ij}

Longitudinal data is data obtained through a repeated observations carried out against a number of objects. Such data have appeared in various fields such as medicine, agriculture, economics. Most longitudinal studies designed to determine the value of the middle of the response as a function of time, with due regard to the role of the explanatory variables.

D. Longitudinal Tobit Analysis

For longitudinal data, tobit models can be defined in the same manner as in the classic longitudinal analysis, the underlying model y^* is a linear mixed models

$$y_{ij}^* \mid b_i = x_{ij}^{'} \beta + z_{ij}^{'} b_i + e_{ij} : N(0, \sigma^2) \tag{4}$$

Where's i refers to the subject i and j .

Estimation of the parameters of the model tobit approximated using Gaussian quadrature. An advantage of the Gaussian quadrature estimate is that the accuracy can be assessed by comparing with a different number of points quadrature. When quadrature obtained by assuming a normal, posterior density estimates of latent variables tend to work well for a continuing response and in large numbers.

E. Glasgow Coma Scale

Glasgow Coma Scale is a standardized scoring system to assess the neurological status of patients with head injury / awareness. GCS accurate value is used to direct treatment and for the prediction of patient outcome. GCS was first introduced in 1970 in Glasgow by Jennet and Teasdale. Rate Glasgow Coma Scale (GCS) is:

Table 1: Glasgow Coma Scale Score

No	Item assessed	Skor
1.	Opened the eyes responses	4 = Spontaneous 3 = Against the verbal stimulus 2 = Against the painful stimulus 1 = No respon
2	Verbal Responses	5 = Oriented on time 4 = Confused 3 = Words irregular 2 = Unclear voice 1 = No voices
3	Motoric responses	6 = obey ordered 5 = Shows the location of pain 4 = Avoid 3 = Abnormal flexion 2 = Abnormal extency 1 = No Respon

F. Stroke

Stroke is a major health problem for today's modern society. The sudden stroke can result in death, physical and mental disability both in their productive age and old age. Stroke is a disorder of neurological dysfunction acute caused by circulatory disorders, and sudden in onset (within seconds) or at least rapidly (within hours) with symptoms and signs according to the focal area of the brain that is impaired. In the case of patients with stroke, a stroke patients often experience trauma / head injury due to falls in the bathroom, fell on the stairs or after waking. Trauma / head injury in stroke patients require ongoing medical records to ensure further medical treatment.

III. RESULTS

A. Data Source

The data used in this study were taken from a longitudinal study of the rehabilitation of stroke patients in which the calculation includes the Glasgow Coma Scale, blood pressure, temperature and pulse rate in patients with stroke. Total sample consisted of 80 patients with stroke in Haji Surabaya hospital with the duration of treatment for 14 days.

B. Research Variable

Response variable (Y) used in this research is the score on the Glasgow Coma Scale (GCS) for the calculation of trauma / injury head stroke patients. GCS' sensor point based on the degree of severity of head injury. GCS' scores can be seen in table 2

Table 2 Respon Variable (Y)

Glasgow Coma Scale	**Score**
Mild Of Head Trauma/Injury	13-15
Middle Of Head Trauma/Injury	9-12
Hard Of Head Trauma/Injury	3-8

The predictor variables were included in this study are blood pressure, body temperature and pulse.

C. Results

In the tabel 3 presented descriptive statistic from the lower margin of Glasgow Coma Scale's score.

Table 3. Ceiling Proportion

GCS	Ceiling Proportion													
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	D13	D14
Mild (13-15)	4%	2%	7%	8%	12%	2%	3%	7%	23%	44%	14%	36%	15%	35%
MIddle (9-12)	7%	4%	7%	5%	12%	4%	3%	14%	23%	5%	21%	27%	12%	25%
Hard (3-8)	0%	0%	3%	3%	5%	4%	5%	8%	3%	2%	12%	12%	2%	23%

Table 3 displays the descriptive statistics (means and standard deviation) and different ceiling proportions across fourteen times. Therefore, different type have substantially different ceiling proportions. The highest score for stroke patiens in the mild group is 44% in day 10 and in the middle group is 27% in day 11. Because different proportions of Glasgow coma scale scores may have different magnitudes of influences on parameter estimates. The sample mean and covariance matrix of the generated data with ceiling threshold = 13 were presented in table 4.

Table 4 The Sammple mean and cov with CT=13

				Ceiling Threshold = 13
	(Y)	X_1	X_2	X_3
Mean vector	7.8	9.2	10.3	12.2
Covariance Matrix				
(Y)	4.3	5.6	5.6	4.7
X_1	4.2	4.8	5.3	3.5
X_2	4.5	3.7	4.8	2.8
X_3	4.3	5.1	3.6	3.0

References

[1] Setyopranoto, I. (2011), *Stroke: Gejala dan Penatalaksanaan*, Fakultas Kedokteran Universitas Gajah Mada

[2] Collin et al. (1988), Barthel Index of Activities of Daily Living

[3] Wang, J., Zhang, Z., McArdle, J.J., dan Salthouse, T.A. (2008), *Investigasing Ceiling Effects in Longitudinal Data Analysis*, University of Notre Dame

[4] Dagne, G.A., Huang, Y. (2011), *Mixed-Effects Tobit Joint Models for Longitudinal data with skewness, detection limits, and measurement errors*, University of South Florida R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[5] Frees, E.W. dan Rosenberg M. (2009), *Session 1a: An Introduction to Two-Part Models and Longitudinal Models for Use in Modeling Health Care Utilization* M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[6] Twisk, J. dan Rijmen, F. (2008), Longitudinal tobit regression : *A new approach to analyze outcome variables with floor or ceiling effects*, VU Medical Centre, Amsterdam, The Netherlands