

# PENERAPAN REGRESI LOGISTIK ORDINAL UNTUK MENGANALISIS TINGKAT KEPARAHAN KORBAN KECELAKAAN LALU LINTAS KABUPATEN BULELENG

Dewa Ayu Made Dwi Yanti Purnami<sup>§1</sup>, I Komang Gde Sukarsa<sup>2</sup>, G. K. Gandhiadi<sup>3</sup>

<sup>1</sup>Jurusan Matematika, Fakultas MIPA - Universitas Udayana [Email: ayudwi\_dewa@yahoo.com]

<sup>2</sup>Jurusan Matematika, Fakultas MIPA - Universitas Udayana [Email: sukarsakomang@yahoo.com]

<sup>3</sup>Jurusan Matematika, Fakultas MIPA - Universitas Udayana [Email: gandhiadigk@yahoo.com]

<sup>§</sup>Corresponding Author

## ABSTRACT

*Ordinal logistic regression is a statistical method for analyzing the response variables that have an ordinal scale consisting of three or more categories. This method is an extension of logistic regression with a binary response variable. In this study the cases studies was the severity of traffic accident victims in Buleleng. The severity of the victims were divided into three categories: minor injuries, serious injuries and died. This research also used six predictor variables, namely age, hours of accident, education, gender, the status of location, and the venicles involved. Models obtained from this study is:*

$$\gamma_1(\underline{X}) = 1.481 - 0.022X_1 - 0.223X_2 - 0.302X_3 + 1.090X_5 \text{ and}$$

$$\gamma_2(\underline{X}) = 2.006 - 0.022X_1 - 0.223X_2 - 0.302X_3 + 1.090X_5$$

*This shows that the variables age, hours of accident, education and the status of location have a significant effect on the severity of traffic accident victims.*

*Keywords: Ordinal logistic regression, severity of traffic accident victims*

## 1. PENDAHULUAN

Analisis regresi merupakan alat statistika yang memanfaatkan hubungan antara dua atau lebih variabel sehingga salah satu variabel bisa diramalkan dari variabel [1]. Variabel yang diramalkan dinyatakan sebagai variabel respon, sedangkan variabel yang meramalkan dinyatakan sebagai variabel prediktor. Analisis regresi dibedakan atas analisis regresi linear dan analisis regresi nonlinear. Salah satu analisis regresi nonlinier adalah analisis regresi logistik. Salah satu metode analisis regresi logistik adalah analisis regresi logistik ordinal. Regresi logistik ordinal merupakan salah satu metode analisis yang dapat digunakan untuk memperoleh hubungan antara variabel respon dengan variabel prediktor. Variabel respon pada regresi logistik ordinal memiliki lebih dari dua kategori yang berskala ordinal dan variabel prediktor berupa data

kategori dan/atau kontinu dengan dua variabel atau lebih ( $k > 2$ ).

Salah satu kasus yang melibatkan variabel respon dengan dua atau lebih kategori yang berskala ordinal adalah tingkat keparahan korban yang mengalami kecelakaan lalu lintas. Berdasarkan penggolongan kecelakaan lalu lintas yaitu kecelakaan lalu lintas ringan, kecelakaan lalu lintas sedang dan kecelakaan lalu lintas berat, maka tingkat keparahan korban kecelakaan lalu lintas dikategorikan menjadi tiga kategori yaitu korban luka ringan, korban luka berat, korban meninggal.

Kecelakaan lalu lintas merupakan hal yang sedang menjadi perhatian khusus. Kecelakaan lalu lintas dinilai menjadi pembunuh terbesar ketiga di Indonesia, di bawah penyakit jantung koroner dan tuberculosis/TBC [2]. Sementara itu angka kecelakaan berdasarkan data Badan Pusat Statistik Bali, pada tahun 2012 kabupaten Buleleng merupakan kabupaten di

Bali yang menempati urutan tertinggi kecelakaan lalu lintas. Melihat tingginya angka kecelakaan yang terjadi di Buleleng penulis ingin menganalisis tingkat keparahan korban kecelakaan lalu lintas di kabupaten Buleleng dengan menerapkan analisis regresi logistik ordinal untuk melihat hubungan antara variabel prediktor dengan tingkat keparahan korban kecelakaan lalu lintas.

Regresi logistik merupakan metode regresi dimana variabel respon  $Y$  dalam bentuk kategorik yaitu variabel biner atau dikotomi [3]. Bentuk umum regresi logistik sebagai berikut [4]:

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p)}{1 + \exp(\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p)}$$

Regresi logistik ordinal merupakan salah satu metode statistika untuk menganalisis variabel respon yang mempunyai skala ordinal yang terdiri atas tiga kategori atau lebih. Variabel prediktor yang dapat disertakan dalam model berupa data kategori atau kontinu yang terdiri atas dua variabel atau lebih [3]. Metode ini merupakan perluasan dari metode regresi logistik dengan variabel respon biner.

Misalkan variabel respon  $Y$  berskala ordinal yang terdiri dari  $K + 1$  nilai dan dinyatakan dengan  $0, 1, 2, \dots, K$ . Ekspresi umum peluang bersyarat  $Y = k$  pada vektor  $x$  dari  $p$  variabel bebas adalah  $P[y = k | x] = \phi_k(x)$ . Misalkan  $\phi_k(x) = \pi_k(x)$ , maka untuk  $K = 2$ , model yang terbentuk dijelaskan oleh persamaan berikut:

$$P(Y = 0 | x) = \frac{e^{g_0(x)}}{e^{g_0(x)} + e^{g_1(x)} + e^{g_2(x)}}$$

$$P(Y = 1 | x) = \frac{e^{g_1(x)}}{e^{g_0(x)} + e^{g_1(x)} + e^{g_2(x)}}$$

$$P(Y = 2 | x) = \frac{e^{g_2(x)}}{e^{g_0(x)} + e^{g_1(x)} + e^{g_2(x)}}$$

Sehingga bentuk persamaan di atas adalah sebagai berikut:

$$P(Y = k | x) = \frac{e^{g_k(x)}}{\sum_{i=0}^K e^{g_i(x)}}$$

dengan  $g_0(x) = 0$  dan  $k = 0, 1, 2$ .

Metode kemungkinan maksimum (*Maximum Likelihood Estimator*) merupakan metode yang digunakan untuk menaksir parameter-parameter model regresi logistik dengan memberikan nilai estimasi  $\beta$  dengan memaksimumkan fungsi *Likelihood* [4].

Berikut merupakan fungsi *likelihood*:

$$l(\beta) = \prod_{i=1}^n [\phi_0(x_i)^{z_{0i}} x \phi_1(x_i)^{z_{1i}} x \dots x \phi_K(x_i)^{z_{Ki}}]$$

$$L(\beta) = \sum_{i=0}^n z_{0i} \ln[\phi_0(x_i)] + z_{1i} \ln[\phi_1(x_i)] + \dots + z_{Ki} \ln[\phi_{Ki}(x_i)]$$

Dimana  $L(\beta) = \ln l(\beta)$  diperoleh dengan mendiferensialkan  $L(\beta)$  terhadap  $\beta$  dan menyamakan dengan nol [3].

Model yang telah diperoleh akan dilakukan pengujian statistik. Pengujian terhadap parameter-parameter model dilakukan untuk mengetahui peran seluruh variabel penjelas baik secara serentak maupun secara parsial.

Interpretasi koefisien dilakukan pada variabel-variabel yang berpengaruh nyata. Pada regresi logistik dengan satu variabel bebas  $\beta_1 = g(x+1) - g(x)$  menunjukkan perubahan nilai logit untuk setiap unit perubahan pada variabel  $x$  [3].

Secara yuridis pengertian korban termasuk dalam pasal 1 undang-undang nomor 13 tahun 2006 tentang perlindungan saksi dan korban yang dinyatakan bahwa korban adalah seorang yang mengalami penderitaan fisik, mental, dan kerugian ekonomi yang diakibatkan oleh suatu tindak pidana [5]. Kecelakaan lalu lintas digolongkan atas :

- a. Kecelakaan lalu lintas ringan;
- b. Kecelakaan lalu lintas sedang; atau
- c. Kecelakaan lalu lintas berat.

## 2. METODE PENELITIAN

Variabel-variabel yang digunakan dalam penelitian ini disesuaikan dengan data sekunder dari laporan Direktorat Lalu Lintas Kepolisian Negara Republik Indonesia kabupaten Buleleng selama bulan januari sampai juni tahun 2014.

Tabel 1. Variabel-variabel Penelitian dan Kategorinya

Variabel	Label	Kategori
Y	Tingkat Keparahan Korban Kecelakaan Lalu Lintas	0 = korban luka ringan 1 = korban luka berat 2 = korban meninggal
X <sub>1</sub>	Usia	(kontinu)
X <sub>2</sub>	Jam Kejadia	1 = pagi 2 = siang 3 = sore 4 = malam
X <sub>3</sub>	Pendidikan	1 = SD 2 = SMP 3 = SMA 4 = Perguruan Tinggi (PT) 5 = Lain-lain
X <sub>4</sub>	Jenis Kelamin	1 = laki-laki 2 = Pprempuan
X <sub>5</sub>	Lokasi Status Jalan	1 = nasional 2 = provinsi 3 = kabupaten/kota 4 = desa

Metode analisis dalam penelitian ini dimulai dengan melakukan analisis statistika deskriptif data untuk mengetahui karakteristik korban kecelakaan lalu lintas Buleleng pada bulan Januari sampai Juni tahun 2014 yang meliputi variabel respon (Y) dan variabel prediktor ( $X_1, X_2, X_3, X_4, X_5, X_6$ ) yang meliputi usia, jam kejadian, pendidikan, jenis kelamin, lokasi status jalan, kendaraan terlibat. Kemudian melakukan analisis regresi logistik ordinal untuk memperoleh faktor-faktor yang memengaruhi tingkat keparahan korban kecelakaan lalu lintas sebagai berikut:

- Melakukan estimasi parameter
- Pembentukan model regresi logistik ordinal dari seluruh variabel prediktor yang berpengaruh dengan variabel respon (Y)
- Melakukan pengujian secara serentak model regresi logistik ordinal
- Melakukan pengujian secara parsial model regresi logistik ordinal
- Apabila terjadi multikolinearitas maka dilakukan analisis regresi logistik bertatar pada langkah (f)

- Melakukan analisis regresi logistik bertatar untuk menyeleksi variabel bebas yang masuk ke dalam model
- Menyatakan model akhir yaitu dengan menggunakan model odd proporsional (*proportional odds model*), modelnya sebagai berikut :

$$c_k(x) = \tau_k - x'\beta$$

$c_k(x)$  : Tingkat keparahan korban kecelakaan lalu lintas yang nilainya dipengaruhi oleh nilai  $x'$ .

$x'$  : Matrik dari variabel prediktor yang dalam hal ini meliputi usia, jam kejadian, pendidikan, jenis kelamin, lokasi status jalan, kendaraan terlibat.

$\beta$  : koefisien variabel prediktor pada model ke-k.

$\tau_k$  : intersep pada model ke-k.

- Melakukan intepretasi terhadap model yang diperoleh.

### 3. HASIL DAN PEMBAHASAN

Data yang diperoleh sebanyak 434 data korban kecelakaan lalu lintas dari bulan Januari sampai dengan Juni tahun 2014. Dari data tersebut terdapat 84.3% yang mengalami luka ringan, 5.5% korban luka berat dan 10.1% korban meninggal dunia.

Selanjutnya pada analisis regresi logistik ordinal, antara 6 variabel prediktor yang diamati dalam penelitian, diketahui model regresi logistik ordinalnya dan ditentukan variabel-variabel prediktor apa saja yang signifikan berpengaruh terhadap tingkat keparahan korban kecelakaan lalu lintas, dengan langkah pertama yaitu pengujian parameter secara serentak.

Pada tahap pengujian parameter secara serentak diperoleh apakah keseluruhan variabel bebas yang digunakan memiliki pengaruh terhadap variabel tak bebas. Adapun hipotesis yang digunakan yaitu:

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

(tidak ada pengaruh sekumpulan variabel prediktor terhadap variabel respon )

$H_1$  : paling sedikit ada satu  $\beta \neq 0$

( terdapat minimal satu dari variabel prediktor berpengaruh terhadap tingkat keparahan korban kecelakaan lalu lintas)

Tabel 2. Hasil *Likelihood Ratio Test*

G	Db	p-value
26.434	6	0.000

Berdasarkan tabel 2 dapat dilihat bahwa pada saat semua variabel prediktor dimasukkan ke dalam model, nilai signifikan ( $p - value$ ) = 0.000 <  $\alpha$ , dibandingkan dengan  $\alpha = 0.05$  (tingkat kepercayaan 95%), maka  $p - value < \alpha$  sehingga keputusannya adalah tolak  $H_0$ . Jadi kesimpulannya adalah terdapat minimal satu variabel prediktor yang berpengaruh terhadap tingkat keparahan korban kecelakaan lalu lintas. Selanjutnya dilakukan pengujian parameter secara parsial.

Hipotesis yang digunakan untuk pengujian parameter secara parsial yaitu:

$H_0 : \beta_k = 0$  (tidak ada pengaruh variabel prediktor yang di uji terhadap variabel respon )

$H_1 : \beta_k \neq 0$  (terdapat pengaruh variabel prediktor yang di uji terhadap variabel respon )

Berdasarkan estimasi parameter maka fungsi logitnya dapat ditulis sebagai berikut:

Tabel 3. Pengujian Parameter secara Parsial

	Koef	SE Coef	Z	p-value
Konstanta 1	1.481	1.506		
Konstanta 2	2.006	1.509		
usia ( $X_1$ )	-0.022	0.007	-2.91	0.004
jam kejadian ( $X_2$ )	-0.223	0.110	-2.03	0.042
pendidikan ( $X_3$ )	-0.302	0.144	-2.09	0.036
jenis kelamin ( $X_4$ )	-0.272	0.307	-0.89	0.375
status jalan ( $X_5$ )	1.090	0.348	3.13	0.002
kendaraan terlibat ( $X_6$ )	0.046	0.386	0.12	0.905

Berdasarkan Tabel 2 dapat dilihat bahwa terdapat empat variabel prediktor yang berpengaruh secara signifikan ( $p - value < 0,05$ ), yaitu variabel usia ( $X_1$ ), jam kejadian ( $X_2$ ), pendidikan ( $X_3$ ) dan lokasi status jalan ( $X_5$ ).

Dari pengujian secara parsial maka model regresi logistik yang diperoleh adalah sebagai berikut:

$$\gamma_1(\underline{X}) = 1.481 - 0.022X_1 - 0.223X_2 - 0.302X_3 + 1.090X_5$$

$$\gamma_2(\underline{X}) = 2.006 - 0.022X_1 - 0.223X_2 - 0.302X_3 + 1.090X_5$$

Pada model regresi logistik, yang digunakan untuk interpretasi koefisien adalah *odds ratio*. Nilai *odds ratio* merupakan rasio antara kecenderungan (resiko) terjadinya suatu peristiwa dalam kelompok kasus dengan kelompok kontrol.

Nilai *odds ratio* masing-masing koefisien pada model terbaik yang diperoleh yaitu model tingkat keparahan korban kecelakaan lalu lintas yang dipengaruhi oleh variabel usia, jam kejadian, pendidikan dan lokasi status jalan adalah sebagai berikut:

Tabel 4. Nilai *Odds Ratio* masing – masing Variabel

Variabel	Koef	<i>Odds ratio</i>
usia ( $X_1$ )	-0.022	0.98
jam kejadian ( $X_2$ )	-0.223	0.80
pendidikan ( $X_3$ )	-0.302	0.74
status jalan ( $X_5$ )	1.090	2.98

Berdasarkan Tabel 3 dapat diinterpretasikan sebagai berikut:

*Odds Ratio*  $X_1$  (usia) = 0.98, artinya setiap kenaikan umur korban kecelakaan lalu lintas 1 tahun, maka menyebabkan perbandingan tingkat keparahan korban kecelakaan lalu lintas dari yang luka ringan ke luka berat atau luka berat ke meninggal dunia menurun sebesar 0.98 kali. Dengan kata lain orang yang memiliki usia lebih besar 1 tahun memiliki peluang meninggal dunia lebih rendah di bandingkan dengan orang yang berusia lebih muda.

*Odds Ratio*  $X_2$  (Jam Kejadian) = 0.80, artinya dengan perubahan jam kejadian dari pagi ke siang atau siang ke sore atau sore ke malam akan mengakibatkan penurunan tingkat keparahan korban kecelakaan lalu lintas dari yang luka ringan ke luka berat atau luka berat ke meninggal dunia adalah sebesar 0.80 kali.

*Odds Ratio*  $X_3$  (Pendidikan) = 0.74, artinya dengan perubahan pendidikan korban dari SD ke SMP atau SMP ke SMA atau SMA ke PT akan mengakibatkan penurunan tingkat keparahan korban kecelakaan lalu lintas dari yang luka ringan ke luka berat atau luka berat ke meninggal dunia adalah sebesar 0.74 kali.

*Odds Ratio*  $X_5$  (lokasi Status Jalan) = 2.98, artinya dengan perubahan lokasi status jalan dari Desa ke Kabupaten atau kabupaten ke provinsi akan mengakibatkan perubahan tingkat keparahan korban kecelakaan lalu lintas dari yang luka ringan ke luka berat atau luka berat ke meninggal dunia adalah sebesar 2.98 kali.

#### 4. KESIMPULAN

Berdasarkan hasil dan pembahasan dalam penelitian ini, maka dapat disimpulkan sebagai berikut:

1. Model regresi logistik ordinal yang terbentuk untuk tingkat keparahan korban kecelakaan lalu lintas adalah sebagai berikut:

$$\gamma_1(\underline{X}) = 1.481 - 0.022X_1 - 0.223X_2 - 0.302X_3 + 1.090X_5$$

$$\gamma_2(\underline{X}) = 2.006 - 0.022X_1 - 0.223X_2 - 0.302X_3 + 1.090X_5$$

2. Faktor-faktor yang berpengaruh terhadap tingkat keparahan korban kecelakaan lalu lintas adalah variabel usia ( $X_1$ ), jam kejadian ( $X_2$ ), pendidikan ( $X_3$ ), dan lokasi status jalan ( $X_5$ ).

Penggunaan metode regresi logistik ordinal pada penelitian ini masih belum cukup signifikan. Hal ini kemungkinan metode yang digunakan tidak sesuai dengan data ataupun mungkin dikarenakan variabel-variabel prediktor yang digunakan dalam penelitian ini masih belum lengkap. Untuk itu pada penelitian berikutnya diharapkan dapat

menggunakan variabel-variabel prediktor yang lebih lengkap.

#### DAFTAR PUSTAKA

- [1] Neter, J., Wasserman, W., & Kutner, M.H. 1997. *Model Linier Terapan BUKU I : Analisis Regresi Linier Sederhana Terjemahan Bambang Sumantri*. Jurusan Statistika FMIPA IPB : Bogor.
- [2] Badan Intelijen Negara., 2013. *Kecelakaan Lalu Lintas: Pembunuh Terbesar Ketiga*. [www.bin.go.id/awas/detil/197/4/21/03/2013/kecelakaan-lalu-lintas-pembunuh-terbesar-ketiga](http://www.bin.go.id/awas/detil/197/4/21/03/2013/kecelakaan-lalu-lintas-pembunuh-terbesar-ketiga). Diakses tanggal 5 September 2014
- [3] Hosmer, D.W and S. Lemeshow. 2000. *Applied Logistic Regression Second Edition*. New York: John Wiley and Sons, Inc.
- [4] Agresti, A., 2002. *Categorical Data Analysis*. 2nd ed. New Jersey: John Wiley and Sons.
- [5] Undang-undang Nomor 13 Tahun 2006 tentang Perlindungan Saksi dan Korban