

REGRESI LOGISTIK UNTUK DATA RESPON POLYTOMOUS

Faula Arina[†]

Jurusan Teknik Industri Universitas Sultan Ageng Tirtayasa Cilegon

Jl. Jend. Sudirman Km. 3 Cilegon, Banten 42435

E-mail: faula.arina71@gmail.com

ABSTRAK

Regresi Logistik untuk data respon polytomous digunakan untuk menentukan hubungan antara peubah respon bersifat lebih dari dua kategori dengan satu atau lebih peubah prediktor kontinyu maupun kategorik. Regresi Logistik untuk data respon polytomous adalah regresi logistik multinomial yang terdiri dari regresi logistik nominal dan ordinal. Pendugaan parameter model regresi logistik nominal dan ordinal dilakukan dengan metode *Maximum Likelihood Estimation* (MLE). Indikator pemilihan model terbaik untuk fungsi MLE adalah *Akaike Information Criteria* (AIC). Berdasarkan analisa data, dapat dibuktikan bahwa peubah respon berskala ordinal selain dianalisis menggunakan regresi logistik ordinal dapat pula dianalisis menggunakan regresi logistik nominal. Hasil pemilihan model regresi ordinal dan nominal terbaik pada kasus hubungan kepuasan konsumen dalam merasakan masakan dengan tambahan keju dari empat merk berdasarkan AIC menghasilkan regresi logistik ordinal sebagai model terbaik.

Kata Kunci: AIC, ordinal, polytomous, regresi logistik, regresi logistik multinomial

[†] Corresponding Author

1. Pendahuluan

Regresi logistik adalah suatu metode untuk menggambarkan hubungan antara peubah respon berkatagori dikotomous atau polytomous dengan satu atau lebih peubah prediktor yang menggunakan skala kategorik maupun interval (Hosmer dan Lemeshow, 1989).

Data polytomous adalah data respon yang mempunyai katagori lebih dari dua misal golongan darah. Peubah respon polytomous yang tidak memiliki urutan disebut peubah nominal sedangkan peubah respon yang memiliki urutan disebut peubah ordinal. Kedua jenis peubah, baik nominal maupun ordinal sering disebut peubah multinomial. Dalam analisa regresi logistik jika peubah respon adalah nominal disebut regresi logistik nominal. Jika peubah respon ordinal disebut regresi logistik ordinal. Regresi logistik multinomial adalah regresi logistik yang peubahrespon berupa data kualitatif berbentuk *multinomial* (nominal atau ordinal). Berdasarkan banyaknya katagori peubah respon, regresi logistik dapat dibedakan menjadi dua yaitu regresi logistik biner dan regresi logistik multinomial. Regresi logistik biner digunakan ketika peubah respon (Y) hanya ada dua kemungkinan misal bekerjadan tidak bekerja. Regresi logistik multinomial dipakai ketika pada peubah respon (Y) terdapat lebih dari dua katagori.

Pendugaan parameter pada model regresi logistik nominal dan ordinal menggunakan metode *Maximum likelihood Estimation* (MLE). Salah satu metode pemilihan model terbaik untk fungsi MLE adalah metode *Akaike Information Creteria* (AIC), dipakai pada contoh berukuran kecil dan jumlah kategori sama.

Tujuan penulisan makalah ini adalah untuk menjelaskan regresi logistik multinomial dan menjelaskan contoh ilustrasi model regresi logistik multinomial pada kasus mengenai pengaruh merk keju terhadap kepuasan konsumen dalam merasakan menu masakan (Mc. Cullagh dan Nelder, 1983, halaman 175) dan akan ditentukan model terbaik antara model logistik nominal dan model regresi logistik ordinal untuk peubah respon berskala ordinal berdasarkan nilai AIC .

2. Metodologi Penelitian

a. Tinjauan Pustaka

a. Regresi Logistik Multinomial

Regresi logistik multinomial adalah salah satu pemodelan yang dapat digunakan untuk mendeskripsikan hubungan beberapa peubah kovariat X dengan suatu peubah respon multinomial (polytomous). Ini termasuk GLM (*Generalized Linier Model*), komponen acak adalah Y bersebaran multinomial (n, π) , komponen sistematik adalah peubah X (dapat diskrit, kontinu atau keduanya) dan fungsi hubungunya adalah *generalized logit* (Dobson, 2001).

Analisis regresi logistik multinomial mempunyai lima pilihan fungsi hubung (*link function*) pada Tabel 1. Penggunaannya tergantung dari sebaran data yang dianalisis. Logit dipakai pada sebagian besar sebaran data, Complementary log log dipakai pada data yang mempunyai kecenderungan bernilai tinggi. Negatif log-log dipakai pada data yang mempunyai kecenderungan bernilai rendah. Probit dipakai jika peubah laten menyebar secara normal, sedangkan chauchit dipakain jika peubah latennya mempunyai nilai yang ekstrim.

Untuk menganalisa peubah respon yang polytomous tergantung dari katagori respon berupa nominal atau ordinal. Jika respon berkatagori nominal, maka digunakan model regresi logistik nominal. Hal ini digunakan ketika tidak ada urutan diantara katagori respon. Jika ada urutan pada katagori respon maka digunakan model regresi logistik ordinal.

Tabel 1. Fungsi Hubung Pada Regresi Logistik Multinomial

<i>Fungsi Hubung</i>	<i>Bentuk Fungsi</i>
Logit	$\log\left(\frac{\pi}{1-\pi}\right)$
Complementary log-log	$\log(-\log(1-\pi))$
Negative log-log	$-\log(-\log(\pi))$
Probit	$\Phi^{-1}(\pi)$
Chauchit	$\text{Tan}(\text{phi}(\pi - 0.5))$

b. Penduga Parameter

Metode yang digunakan untuk mengestimasi parameter model regresi logistik multinomial pada penulisan ini adalah MLE. Metode MLE dapat dilakukan jika antar pengamatan diasumsikan bebas. Penyelesaian menggunakan metode MLE akan bergantung dari jumlah katagori respon yang digunakan. Metode MLE pada persamaan regresi logistik biner memakai asumsi mengikuti sebaran binomial sedangkan pada regresi logistik politomus diasumsikan mengikuti sebaran multinomial.

Pengamatan y_1, \dots, y_n yang independen dan $y_i = (y_{i1}, y_{i2}, \dots, y_{ik})$, $\sum_j y_{ij} = m_i$ dan probabilitas peubah respon π_1, \dots, π_n , $\sum_j \pi_{ij} = 1$.

Fungsi kemungkinan dinyatakan sebagai berikut

$$L(\pi; y) = \prod_j [\pi_{ij}^{y_{ij}}]$$

Sedangkan fungsi log kemungkinannya adalah

$$l(\pi; y) = \sum_{ij} y_{ij} \log \pi_{ij}$$

Langkah berikutnya adalah memperoleh penduga parameter dari regresi logistik ordinal. Yaitu memaksimalkan fungsi log kemungkinan terhadap parameternya

$$\frac{\partial l(\boldsymbol{\pi}; \mathbf{y})}{\partial \pi_{ij}} = \frac{y_{ij} - m_i \pi_{ij}}{\pi_{ij}}$$

notasi dalam bentuk matriks menjadi

$$\begin{aligned} \frac{\partial l(\boldsymbol{\pi}; \mathbf{y})}{\partial \boldsymbol{\pi}_i} &= m_i \boldsymbol{\Sigma}_i^{-1} (\mathbf{y}_i - m_i \boldsymbol{\pi}_i) \\ &= m_i \boldsymbol{\Sigma}_i^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_i) \end{aligned}$$

bila ada n vektor dikelompokkan dalam satu matriks

$$\frac{\partial l(\boldsymbol{\pi}; \mathbf{y})}{\partial \boldsymbol{\pi}} = \mathbf{M} \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\mu})$$

dengan

$$\begin{aligned} \mathbf{M} &\text{ matriks diagonal } n \times n \\ \boldsymbol{\Sigma} &= \text{diag} \{ \boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_n \} \end{aligned}$$

Jika \mathbf{z}_i = Vektor kumulatif respon

$\boldsymbol{\gamma}_i$ = Vektor probabilitas kumulatif

$$\begin{aligned} \frac{\partial l(\boldsymbol{\gamma}; \mathbf{y})}{\partial \boldsymbol{\gamma}_i} &= m_i \boldsymbol{\Gamma}_i^{-1} (\mathbf{z}_i - m_i \boldsymbol{\gamma}_i) \\ \frac{\partial l(\boldsymbol{\pi}; \mathbf{y})}{\partial \gamma_{ij}} &= \frac{\partial l(\boldsymbol{\pi}; \mathbf{y})}{\partial \pi_{ij}} - \frac{\partial l(\boldsymbol{\pi}; \mathbf{y})}{\partial \pi_{i, j-1}} \end{aligned} \text{ untuk}$$

$1 < j < k$

$$\begin{aligned} \frac{\partial l(\boldsymbol{\pi}; \mathbf{y})}{\partial \alpha_r^*} &= \sum_{ij} \frac{\partial l(\boldsymbol{\pi}; \mathbf{y})}{\partial \gamma_{ij}} \frac{\partial \gamma_{ij}}{\partial \alpha_r^*} \\ &= \sum_{ij} x_{ijr}^* \partial \gamma_{ij} (1 - \partial \gamma_{ij}) \frac{\partial l(\boldsymbol{\pi}; \mathbf{y})}{\partial \gamma_{ij}} \\ \frac{\partial l(\boldsymbol{\pi}; \mathbf{y})}{\partial \gamma_{ij}} &= \frac{y_{ij} - m_i \pi_{ij}}{\pi_{ij}} - \frac{y_{i, j-1} - m_i \pi_{i, j-1}}{\pi_{i, j-1}} \end{aligned}$$

untuk $i < j < k$

Persamaan *likelihood* pada regresi logistik multinomial adalah persamaan non linear dalam parameter, sehingga untuk menyelesaikan persamaan tersebut sampai diperoleh nilai estimasi parameternya digunakan algoritma Newton Raphson.

$$\boldsymbol{\alpha}^{(1)} = \boldsymbol{\alpha}^{(0)} + \left[-l''(\boldsymbol{\alpha}^{(0)}) \right]^{-1} l'(\boldsymbol{\alpha}^{(0)})$$

dengan

$$l'(\boldsymbol{\alpha}^{(0)}) = \mathbf{X}^T (\mathbf{y} - \boldsymbol{\mu})$$

$$l''(\boldsymbol{\alpha}) = \mathbf{X}^T \mathbf{W} \mathbf{X}$$

sehingga

$$\boldsymbol{\alpha}^{(1)} = \boldsymbol{\alpha}^{(0)} + \left[\mathbf{X}^T \mathbf{W} \mathbf{X} \right]^{-1} \mathbf{X}^T (\mathbf{y} - \boldsymbol{\mu})$$

c. Interpretasi Koefisien Regresi Logistik Pada Data Respon Polytomous

Interpretasi untuk model regresi logistik menggunakan nilai rasio odds. Misal untuk peubah X yang berskala nominal (x_1 dan x_2) rasio odds pada katagori $Y \leq s$ merupakan perbandingan antara x_1 dan x_2 yang persamaannya sebagai berikut (Agresti 2007).

$$\begin{aligned} L_s(x_1) - L_s(x_2) &= \log \left(\frac{P(Y \leq s | x_1 / Y > s / x_1)}{P(Y \leq s | x_2 / Y > s / x_2)} \right) \\ &= \alpha_i (x_1 - x_2) \end{aligned}$$

dengan $i=1,2,\dots,p$ (p merupakan banyaknya parameter dan $s = 1,2,\dots, s-1$). Parameter α_i menunjukkan perubahan nilai fungsi logit yang disebabkan oleh perubahan satu unit peubah penjelas ke satu disebut log odds, misal antara x_1 dan x_2 yang persamaannya sebagai berikut

$$\text{Log} (\psi [x_1, x_2]) = \alpha_i (x_1 - x_2)$$

sehingga diperoleh penduga untuk rasio odds ($\hat{\psi}$)

$$\hat{\psi} = \exp [\alpha_i (x_1 - x_2)]$$

Untuk peubah penjelas katagorik jika rasio odds bernilai > 1 maka odds saat x_1 lebih besar daripada saat x_2 atau dengan kata lain $P(Y \leq s | x_1)$ akan selalu lebih besar dari $P(Y > s | x_1)$. Sehingga dapat dinyatakan saat x_1 peluang untuk $Y \leq s$ lebih besar daripada saat x_2 .

Untuk peubah penjelas x berskala kontinu, odds saat x mengalami kenaikan 1 unit adalah sebesar $\exp [\alpha_i (x_1 - x_2)]$ kali odds saat x belum mengalami kenaikan. Jika nilai rasio bernilai > 1 maka peluang untuk $Y \leq s$ pada saat x mengalami kenaikan lebih besar dari pada saat x belum mengalami kenaikan. Sedangkan pada peubah kontinu berskala besar diperlukan perubahan unit sebesar c untuk interpretasinya dengan rasio odds sebesar $\exp [c \alpha]$.

d. Uji Kebaikan Model

Ada beberapa kriteria untuk mengukur kebaikan model regresi logistik diantaranya adalah Pearson Chi square residual, Uji Rasio Likelihood dan AIC.

Kriteria pertama untuk mengukur kebaikan model adalah nilai Pearson Chi Square residual dengan persamaan

$$r_i = \frac{o_i - e_i}{\sqrt{e_i}}$$

dimana o_i dan e_i adalah nilai pengamatan dan nilai

harapan untuk setiap $i = 1,2,\dots,n$. Nilai $\sum_{i=1}^n r_i$

mengikuti sebaran χ^2 dengan derajat bebas $n - p$, dimana p adalah banyaknya parameter.

Pearson Chi square dengan derajat bebas np maka model yang dihasilkan kemungkinan mempunyai tingkat kesesuaian yang cukup.

Kriteria kedua adalah uji rasio likelihood yang dapat diimplimentasikan untuk menduga kesesuaian dari pendugaan parameter regresi dengan menggunakan MLE. Uji ini untuk melihat kontribusi peubahpenjelas terhadap peubahrespon di dalam model. Dengan demikian uji hipotesis yang dilakukan adalah

$$H_0 = \alpha_0 = \alpha_1 = \dots = \alpha_p$$

$$H_1 = \text{sedikitnya ada 1 dari } \alpha_p \text{ yang tidak sama}$$

Statistik uji yang digunakan $G = 2(l_1 - l_0)$ dimana

l_1 adalah likelihood tanpa peubahpenjelas dan l_0 adalah likelihood dengan peubahpenjelas. Nilai G mengikuti sebaran χ^2 dengan bebas k .

Kriteria ketiga untuk mengukur kebaikan model adalah AIC (Akaike Information Criteria) dan SC (Schwartz Criteria). AIC didefinisikan sebagai $G^2 - 2d$, d adalah derajat bebas. Nilai AIC yang semakin kecil mengidentifikasikan model yang baik.

b. Metode Penelitian

a. Data penelitian

Penelitian ini menggunakan data pada Buku Mc Culagh dan Nelden (1983) halaman 175. Hasil penelitian Dr. Graeme Newell dari Hawkesbury Agricultural College yang melakukan eksperimen dimana pengunjung restoran sebanyak 52 orang diminta merasakan 4 merk Keju A,B,C dan D yang ditambahkan pada menu makanannya. Untuk mengetahui keju mana yang kualitas bagus menurut konsumen dan hubungan antara kepuasan pengunjung dengan penambahan keju pada menu makanannya. Data penelitian di Tabel 2.

Tabel 2. Data Penelitian

Keju	Kategori respon									Total
	1	2	3	4	5	6	7	8	9	
A	0	0	1	7	8	8	19	8	1	52
B	6	9	12	11	7	6	1	0	0	52
C	1	1	6	8	23	7	5	1	0	52
D	0	0	0	1	3	7	14	16	11	52
Total	7	10	19	27	41	28	39	25	12	208

1 = sangat tidak enak 9 = sangat enak

b. Prosedur Analisis Data

1. Data di analisa menggunakan program R 3.5.3 dengan menggunakan analisa regresi logistik ordinal untuk melihat hubungan antara tingkat kepuasan dengan merk keju.
2. Data dimodelkan kembali menggunakan analisa regresi logistik nominal.

3. Hasilnya dibandingkan berdasarkan nilai AIC. Nilai AIC yang paling kecil sebagai model yang paling baik.

3. Hasil dan Pembahasan

a Analisa Regresi Logistik Ordinal

Data pada Tabel 2 diolah menggunakan Regresi Logistik Ordinal dengan *Proportional odds Model* dipilih dikarenakan kategori respon dari tidak enak ke enak sekali (9 kategori) sebagai skala ordinal dan pada *Proportional odds Model* ini perbandingan dilakukan antara peluang mendapatkan nilai respon sama dengan atau kurang dari nilai tertentu dengan peluang mendapatkan nilai respon yang lebih besar. Dengan fungsi hubung logit modelnya sebagai berikut

$$\log \left(\frac{\pi_1(x) + \dots + \pi_j(x)}{\pi_{j+1}(x) + \dots + \pi_k(x)} \right) =$$

$$\alpha_{0j} + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 \quad \text{untuk } j = 1, 2, \dots, 8$$

$$k = 1, 2, \dots, 9$$

dengan

$$x_1 = \begin{cases} 1, & \text{keju B} \\ 0, & \text{lainnya} \end{cases}$$

$$x_2 = \begin{cases} 1, & \text{keju C} \\ 0, & \text{lainnya} \end{cases}$$

$$x_3 = \begin{cases} 1, & \text{keju D} \\ 0, & \text{lainnya} \end{cases}$$

Model Logistik Ordinal yang digunakan dalam analisa data ada 8 diantaranya

$$\log \left(\frac{P(y \leq 1)}{P(y > 1)} \right) = \log \left(\frac{\pi_1}{\pi_2 + \dots + \pi_9} \right) =$$

$$\alpha_{01} + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3$$

Selanjutnya digunakan Program R 3.5.3 seperti pada Tabel 3 untuk menentukan estimasi parameter menggunakan regresi logistik ordinal dengan model *proportional odd* di Program R digunakan `polr()` dan untuk menentukan estimasi parameter regresi logistik nominal dengan menggunakan `Generalize logit` : multinom (Crawly, 2013)

Tabel 3. Program R untuk Regresi Logistik Ordinal dan Regresi Nominal

```
cheese <-read.table("chees.dat",
col.names=c("Cheese", "response", "count"))
```

```
is.factor(cheese$response)

#### make it into an ordered factor
cheese$response<-factor(cheese$response,
ordered=T)
library(MASS)
che.polr <- polr(response ~ cheese, data=cheese,
weight=count, method="logistic")
summary(che.polr)
Anova(che.polr) #Type II test
library(nnet)
mod.multinom<-multinom(response~cheese,
data=cheese, weight=count)
summary(mod.multinom, Wald=TRUE)
Anova(mod.multinom)
```

Estimasi parameter pada regresi Logistik Ordinal pada Tabel 4. Estimasi parameter signifikan sebab nilai $T > |1.96|$, kecuali intersep 6 tidak signifikan karena kurang dari 1.96 atau $0.1673 < |1.96|$. Kualitas keju menurut konsumen adalah Keju $D > A > C > B$ hal ini dapat dilihat dari nilai dugaan/koeffisien D positif (1.613), A(nol), C negatif (-1.710) dan B negatif (-3.352). Keju B dibanding keju A kan menurut preferensi konsumen adalah berdasar Estimasi odds ratio $\exp(-3.352) = 0.035$. The standard error = 0.4287 dan t-value = -7.819 with df=1, menunjukkan signifikan bahwa preference keju B terhadap A lebih jelek.

Tabel 4. Nilai Dugaan Parameter pada Regresi Logistik Ordinal

Parameter	df	Dugaan	Galat baku	T
Intersep 1	1	-5.47	0.52	-10.44
Intersep 2	1	-4.41	0.48	-10.31
Intersep 3	1	-3.31	0.37	-8.95
Intersep 4	1	-2.24	0.32	-6.87
Intersep 5	1	-0.91	0.28	-3.20
Intersep 6	1	0.04	0.26	0.17
Intersep 7	1	1.55	0.30	5.12
Intersep 8	1	3.10	0.40	7.65
Keju B	1	-3.35	0.43	-7.82
Keju C	1	-1.71	0.37	-4.60
Keju D	1	1.61	0.38	4.24

Tabel 5 menunjukkan bahwa merk keju berpengaruh terhadap kepuasan konsumen dalam

menikmati menu makanan karena nilai LR Chisq 148.45 dengan signifikansi < 0.05 .

b. Regresi Logistik Nominal

Data pada Tabel 2 di modelkan kembali dengan menggunakan regresi logistik Nominal yang peubahresponnya ada 9 kategori. Satu katagori diantaranya dipilih sebagai katagori acuan dan katagori 1 dipilih sebagai acuan. Model logit untuk katagori selain katagori acuan dengan fungsi hubung logit, modelnya sebagai berikut

$$\text{logit}(\pi_j(x)) = \log\left(\frac{\pi_j(x)}{\pi_1(x)}\right) =$$

$$\alpha_{0j} + \alpha_{1j}x_1 + \alpha_{2j}x_2 + \alpha_{3j}x_3 \text{ untuk } j = 2, \dots, 9$$

dengan

$$x_1 = \begin{cases} 1, & \text{keju B} \\ 0, & \text{lainnya} \end{cases}$$

$$x_2 = \begin{cases} 1, & \text{keju C} \\ 0, & \text{lainnya} \end{cases}$$

$$x_3 = \begin{cases} 1, & \text{keju D} \\ 0, & \text{lainnya} \end{cases}$$

Salah satu model dari 8 model regresi logistik nominal adalah

$$\log\left(\frac{\pi_2}{\pi_1}\right) = -13.507 + 13.913x_1 + 13.507x_2 + 6.866x_3$$

$$\pi_1(x) = \frac{e^{-13.507+13.913x_1+13.507x_2+6.866x_3}}{1 + e^{-13.507+13.913x_1+13.507x_2+6.866x_3}}$$

Interpretasi Logit 1 Menunjukkan tingkatan kepuasan konsumen kategori 2 dengan kategori 1 (tidak enak) sebagai kategori pembanding mempunyai model logit 1

Tabel 5. Anova Pada Regresi Logistik Ordinal

LR Chisq	Df	Pr(>Chisq)
keju	148.45	3 < 2.2e-16 ***

Tabel 6 menunjukkan bahwa merk keju berpengaruh terhadap kepuasan konsumen dalam menikmati menu makanan karena nilai LR Chisq 148.45 dengan signifikansi < 0.05 .

Tabel 6 Anova Pada Regresi Logistik Nominal

	LR Chisq	Df	Pr(>Chisq)
Keju	168.76	24	< 2.2e-16 ***

McCullagh, P. and Nelder, J.A. (1983). *Generalized Linear Models Second Edition*. Chapman and Hall.

c. Penentuan Model Regresi Logistik Terbaik

Tabel 7 menunjukkan nilai AIC regresi logistik ordinal sebesar 773.3479 lebih kecil dari pada AIC regresi logistik Nominal sebesar 755.0397 sehingga model regresi logistik ordinal lebih baik dari pada model regresi logistik nominal dalam kasus mengenai hubungan merk keju dengan tingkat kepuasan konsumen adalah merk keju berhubungan dengan tingkat kepuasan konsumen.

Tabel 7 AIC Regresi Logistik

Kriteria	Regresi Logistik	
	Ordinal	Nominal
AIC	733.3479	755.0397

4. Kesimpulan

1. Aplikasi regresi logistik ordinal dengan kategori peubah respon terurut (ordinal) pada suatu kasus mengenai hubungan antara tingkat kepuasan konsumen dengan merk keju adalah merk keju berhubungan dengan tingkat kepuasan konsumen dengan AIC= 733.3479.
2. Aplikasi regresi logistik nominal dengan kategori peubah respon terurut (ordinal) pada suatu kasus mengenai hubungan antara tingkat kepuasan konsumen dengan merk keju adalah merk keju berhubungan dengan tingkat kepuasan konsumen dengan AIC=755.0397.
3. Berdasarkan nilai AIC diperoleh model yang terbaik pada kasus mengenai hubungan antara tingkat kepuasan konsumen dengan merk keju adalah regresi logistik ordinal dengan AIC = 733.3479.

DAFTAR PUSTAKA

- Agresti, A. (2007). *An Introduction to Categorical Data Analysis Second Edition*. John Wiley and Sons, Inc.
- Crawly, M.J. (2013). *The R Book Second Edition*. Chapman and Hall. John Wiley and Sons, Inc.
- Dobson, A.J. (2001). *An Introduction to Generalized Linear Models Second Edition*. Chapman Hall/CRC Texts in Statistical Science Series.
- Hosmer, D.W dan Lemeshow S., (1989), *Applied Logistic Regression*, John Wiley and Sons, New York.