



e-ISSN 2657-232X

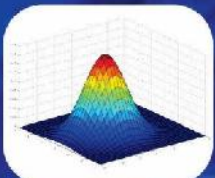
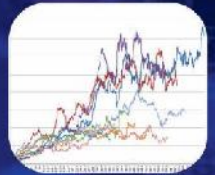
Terbitan II

Mei 2022

# PROSIDING

SEMINAR NASIONAL  
MATEMATIKA, STATISTIKA,  
DAN APLIKASINYA

*“MATHEMATICS AND STATISTICS  
FOR SUPPORTING RESEARCH  
IN THE NEW NORMAL ERA”*



Jurusan Matematika  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Universitas Mulawarman

# PROSIDING

## SEMINAR NASIONAL MATEMATIKA, STATISTIKA, DAN APLIKASINYA

Jurusan Matematika  
Fakultas Matematika dan Ilmu Pengetahuan Alam,  
Universitas Mulawarman, Samarinda, Kalimantan Timur

Ruang lingkup:

- √ Matematika
- √ Statistika
- √ Pendidikan Matematika
- √ Statistika Terapan
- √ Matematika Terapan

Penerbit:  
Jurusan Matematika  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Universitas Mulawarman  
Jalan Barong Tongkok No. 4 Kampus Gunung Kelua  
Samarinda Kalimantan Timur 75123

# PROSIDING

## SEMINAR NASIONAL MATEMATIKA, STATISTIKA, DAN APLIKASINYA

Terbitan II, Mei 2022

Fakultas Matematika dan Ilmu Pengetahuan Alam,  
Universitas Mulawarman, Samarinda, Kalimantan Timur

### Reviewer:

Dr. Sri Wahyuningsih, M.Si.	Rito Goejantoro, S.Si. M.Si
Dr. Suyitno, S.Pd., M.Sc	Yuki Novia Nasution, S.Si., M.Si
Dr. Syaripuddin, S.Si., M.Si	Wasono, S.Si., M.Si
Dr. Darnah Andi Nohe, M.Si	Ika Purnamasari, S.Si., M.Si
Dr. Sifriyani, M.Si	Memi Nor Hayati, S.Si, M.Si
Dr. Fathurahman, M.Si	Meiliyani Siringoringo, S.Si., M.Si

### Ketua:

Dr. Darnah Andi Nohe, S.Si., M.Si

### Editor:

Qonita Qurrota A'yun. S.Si., M.Sc  
Fidia Deny Tisna Amijaya, M.Si  
Surya Prangga, M.Si

### Desain Cover:

Moh. Nurul Huda, M.Si

**SUSUNAN PANITIA**

**Dewan Kehormatan**

Prof. Dr. H. Masjaya, M.Si

Vika Kurnia Lestari, S.Si

Kurniawati, S.Si

Fedy Harlanto, S.Si

**Pengarah**

Dr. Eng. Idris Mandang, M.Si

**Seksi Acara**

Wasono, S.Si. M.Si

Dr. Sri Wahyuningsih, M.Si

**Penanggung Jawab**

Wakil Dekan Bidang Akademik,  
Kemahasiswaan, dan Alumni

Dr. Sifriyani, M.Si

Memi Nor Hayati, S.Si., M.Si

Dr. Yanti Puspitasari, M.Si

Ruli Yuniarto, S.Si

Muhammad Rasyid Rahman, S.Stat

**Ketua**

Dr. Darnah Andi Nohe. M.Si

**Seksi Publikasi**

Dr. Suyitno, S.Pd., M.Sc

**Wakil Ketua**

Dr. M. Fathurahman, M.Si

Rito Goejantoro, S.Si., M.Si

Fidia Deny Tisna Amijaya, M.Si

Dr. Syaripuddin, M.Si

Yuki Novia Nasution, S.Si., M.Sc

Surya Prangga, S.Si., M.Si

Moh. Nurul Huda, S.Si, M.Si

**Seksi Kesekretariatan**

Meiliyani Siringoringo, S.Si., M.Si

Qonita Qurrota A'yun, S.Si., M.Sc

Ika Purnamasari, S.Si., M.Si

Ryan Rahmad Ramadhan, S.Si

Nurul Huda, SE

## **KATA PENGANTAR**

Segala puji bagi Allah SWT karena atas rahmat-Nya sehingga Prosiding Seminar Nasional Matematika, Statistika, dan Aplikasinya dapat diterbitkan. Seminar Nasional Matematika, Statistika, dan Aplikasinya diselenggarakan oleh Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA), Universitas Mulawarman Kalimantan Timur pada masa pandemi *Coronavirus Disease 2019* (COVID-19) sehingga dilaksanakan dalam jaringan (daring) menggunakan aplikasi *Zoom Meeting*. Adapun tema kegiatan seminar ini adalah “*Mathematics and Statistics for Supporting Research in the New Normal Era*”.

Munculnya COVID-19 di Indonesia pada tanggal 2 Maret 2020 menyebabkan pemerintah Indonesia mengambil kebijakan penerapan *social distancing* pada tanggal 15 Maret 2020 untuk menekan penyebaran virus mematikan tersebut. Pada bulan Juni 2020, pemerintah Indonesia menerapkan *new normal*, yaitu melakukan aktivitas seperti kehidupan pada umumnya namun tetap memperhatikan protokol kesehatan. Meskipun berada pada era *new normal*, antusias peneliti untuk melaksanakan kegiatan penelitian tetap tinggi.

Berbagai penelitian di bidang Matematika dan Statistika terus dilakukan, Pada bidang Matematika, dilakukan pemodelan penyebaran COVID-19 pada daerah tertentu, membantu memahami dan mengidentifikasi hubungan penyebaran COVID-19 dengan berbagai parameter epidemiologi, membantu dalam perencanaan masa depan dan mempertimbangkan langkah-langkah pengendalian yang tepat. Berbagai penelitian lain telah menjadikan matematika sebagai *basic of science* untuk menemukan hal-hal baru dalam pengembangan ilmu pengetahuan dan teknologi. Pada bidang Statistika, kasus COVID-19 dapat dideskripsikan dengan baik, mengetahui fluktuasi kasus yang terjadi, mengidentifikasi faktor-faktor yang berpengaruh terhadap penyebaran COVID-19, dan memprediksi kasus COVID-19 beberapa periode yang akan datang.

Kegiatan seminar ini menjadi sarana bertukar informasi dan diskusi mengenai penelitian-penelitian terkini dan ilmu pengetahuan yang sedang berkembang baik berkaitan dengan pandemi COVID-19 maupun hal lainnya. Selain itu, kegiatan seminar ini memberikan motivasi dan gagasan baru kepada peserta untuk

melakukan penelitian lain terkait matematika, statistika, dan aplikasinya serta menghasilkan prosiding. Prosiding yang telah tersusun ini berisi tentang penelitian-penelitian terbaru terkait dengan bidang matematika dan statistika, serta aplikasinya dan telah disajikan oleh pemateri, baik praktisi dari berbagai instansi, mahasiswa, maupun dosen dari perguruan tinggi diseluruh Indonesia. Prosiding yang telah tersusun ini dapat diselesaikan dengan baik atas dukungan dari berbagai pihak sehingga kami mengucapkan terima kasih dan apresiasi yang setinggi-tingginya kepada:

1. Rektor Universitas Mulawarman.
2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Mulawarman beserta jajarannya.
3. Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Mulawarman.
4. *Editor* dan *Reviewer* yang telah memberikan penilaian dan seleksi terhadap makalah-makalah prosiding.
5. Panitia Pelaksana Seminar Nasional Matematika, Statistika, dan Aplikasinya tahun 2021.
6. Seluruh peserta seminar, baik praktisi dari berbagai instansi maupun civitas akademika dari berbagai perguruan tinggi di seluruh Indonesia yang telah berkontribusi dalam pelaksanaan kegiatan ini.

Perbaikan dan peningkatan kualitas Prosiding Seminar Nasional Matematika, Statistika, dan Aplikasinya menjadi upaya yang terus dilakukan sehingga kritik yang konstruktif menjadi input yang sangat kami harapkan. Akhir kata, semoga Prosiding Seminar Nasional Matematika, Statistika, dan Aplikasinya dapat memberikan manfaat yang sebesar-besarnya bagi kita semua.

Samarinda, Mei 2022

Penyusun

## DAFTAR ISI

Halaman Judul.....	i
Reviewer dan Editor.....	ii
Panitia.....	iii
Kata Pengantar .....	v
Daftar Isi.....	vii
<b>Makalah</b>	
Optimalisasi Waktu Penjadwalan Dengan Program Evaluation and <i>Review Technique</i> (PERT) Untuk Pembangunan Rumah Tipe 36 Pada PT. Arisko Di Sambutan Samarinda <i>Indah Fatikawati, Langgeng Prasadewo S.A.W.B., Safiah</i> .....	1
Implementasi <i>Game Theory</i> Pada Strategi Pemasaran Ban Roda Dua di Kecamatan Sambutan Samarinda <i>Benny Maryam Halim, Lisnawati , dan Nuryatma</i> .....	9
Analisis Faktor-Faktor yang Memengaruhi Jumlah Penderita Tuberkulosis Menggunakan <i>Geographically Weighted Regression</i> di Provinsi Jawa Timur <i>Angeliya Milla Khariyani, Kismiantini, Ezra Putranda Setiawan</i> .....	16
Identifikasi Faktor-Faktor yang Memengaruhi Indeks Pembangunan Manusia (IPM) di Kalimantan Barat Tahun 2020 Menggunakan <i>Geographically Weighted Regression</i> <i>Andriati Langiran, Kismiantini, Ezra Putranda Setiawan</i> .....	27
Pemodelan Prevalensi Gizi Buruk Pada Balita di Kalimantan Tengah Menggunakan <i>Geographically Weighted Regression</i> <i>Ryan Sidiq Permana, Kismiantini, Ezra Putranda Setiawan</i> .....	36
Model <i>Geographically Weighted Weibull Regression</i> dengan Kriteria Penentuan <i>Bandwidth Optimum Akaike Information Criterion</i> (Studi Kasus: Indikator Pencemaran Air Biochemical Oxygen Demand di Daerah Hutan Tropis Lembab Das Mahakam Tahun 2019) <i>Novianti, Suyitno, Meiliyani Siringoringo</i> .....	43
Peramalan Indeks Harga Konsumen Provinsi Kalimantan Timur dengan Metode <i>Grey Double Exponential Smoothing Holt</i>	



<i>Galuh Batul Nabilah, Yuki Novia Nasution, Ika Purnamasari.....</i>	69
Perbandingan Model Chen dan Model Lee Pada Metode <i>Fuzzy Time Series</i> Untuk Peramalan Produksi Kelapa Sawit Provinsi Kalimantan Timur	
<i>Ipan, Syaripuddin, Darnah Andi Nohe.....</i>	81
Peramalan Data Ekspor Kalimantan Barat dengan Metode <i>Autoregressive Integrated Moving Average (ARIMA)</i>	
<i>Ahmad Aliful Abidin, Paulus Felipe Buiney, Darnah Andi Nohe.....</i>	96
Regresi Data Panel untuk Memodelkan Persentase Kemiskinan di Kalimantan Timur	
<i>Muhammad Taufik Nur, Deva Khoirotunnisa, Wiwit Widyaningsih, Darnah Andi Nohe .....</i>	108
Analisis Faktor yang Berpengaruh Terhadap Tingkat Kemiskinan di Provinsi Kalimantan Timur Menggunakan Regresi Data Panel	
<i>Edy Widodo, Dhea Laksmi Arsyia Primananda, Diannita Eka Putri, Falah Novayanda Adlin, Shafa Bunga Faradilla.....</i>	122
Pemilihan Model Regresi Bernoulli Terbaik	
<i>M. Fathurahman.....</i>	138
Pemodelan Regresi Data Panel Pada Faktor yang Mempengaruhi Tingkat Pengangguran Terbuka di Jawa Barat	
<i>Muhammad Irfan Rizki, Farhat Gumelar, Jessica Jesslyn Cerelia, Teguh Ammar, Adhiyatma Nugraha.....</i>	147
Analisa Ketepatan Waktu Lulus Mahasiswa FKIP Universitas Islam Malang Berdasarkan Penguasaan Kompetensi Menggunakan Klasifikasi CHAID	
<i>Siti Nurul Hasana, Nuse Aliyah Rahmati.....</i>	162
Analisis Faktor-Faktor yang Memengaruhi Produk Domestik Regional Bruto (PDRB) di Kalimantan Timur dengan Regresi Data Panel	
<i>Anisa Yuliana, Muhammad S. Basri, Siti N. Rahmah, Darnah A. Nohe.....</i>	177
Pemodelan Tingkat Pengangguran Terbuka Di Pulau Kalimantan Dengan Regresi Data Panel	
<i>Nuning Kusumaningrum, Jordan Nata Permana, Khairunnisa, Darnah Andi Nohe.....</i>	196



Analisis Regresi Data Panel Pada Faktor yang Memengaruhi Persentase Kemiskinan di Kalimantan Timur	
<i>Rina Aprilianti, Gerald C. Messakh, Sinta Nur Asiah, Darnah A. Nohe.....</i>	211
Pemodelan Indeks Kedalaman Kemiskinan di Provinsi Kalimantan Timur dengan Regresi Data Panel	
<i>Ni Made Shantia A., Anggi J. Utami, Rani Mirnawati, Darnah Andi Nohe....</i>	224
Regresi Data Panel Pada Ketimpangan Pendapatan Daerah di Provinsi Kalimantan Timur	
<i>Nur Afifah Salsabila, Handy K. Juliarto, Al Fitri Syawal, Darnah A. Nohe...</i>	241
Regresi Data Panel Untuk Mengetahui Faktor yang Memengaruhi Jumlah Penduduk Miskin di Kalimantan Timur	
<i>Iqbal Firman Alamsyah, Rut Esra, Salwa Awalia, Darnah Andi Nohe.....</i>	254
Hubungan Jumlah Penduduk Miskin dengan Berat Badan Lahir Rendah di Kalimantan Timur Menggunakan Korelasi Pearson dan Spearman	
<i>Andreas Sergio, Muhammad A. Zen, Risa K. Wahyuni, Darnah A. Nohe.....</i>	267
Studi Simulasi dan Aplikasi: Estimator Deret Fourier Pada Pemodelan Regresi Nonparametrik	
<i>Andrea Tri Rian Dani, Ardiana Fatma Dewi, Ludia Ni'matuzzahroh.....</i>	279
Pemodelan Rata-Rata Lama Sekolah di Provinsi Nusa Tenggara Timur (NTT) Menggunakan Pendekatan Regresi Nonparametrik <i>Spline Truncated</i>	
<i>Ludia Ni'matuzzahroh, Andrea Tri Rian Dani.....</i>	289
Pemodelan Gizi Buruk Pada Balita di Provinsi Jawa Timur Menggunakan <i>Geographically Weighted Regression (GWR)</i>	
<i>Qatrunnada Azkia, Kismiantini, Ezra Putranda Setiawan.....</i>	302
Pemodelan <i>Spatial Autoregressive Quantile Regression</i> dalam Faktor yang Mempengaruhi Tingkat Incident Rate Demam Berdarah <i>Dengue</i> di Jawa Barat	
<i>Muhammad Irfan Rizki, Teguh Ammar.....</i>	312
Perbandingan Tingkat Konsistensi Uji Distribusi Normalitas Pada Data Tingkat Pengangguran di Jawa	
<i>Ineu Sintia, Muhammad Danil Pasarella, Darnah Andi Nohe.....</i>	322

Penentuan Ketepatan Klasifikasi Indeks Kedalaman Kemiskinan di Indonesia dengan Model Logit

*Deviyana Nurmin, Lisa D. N. Khasanah, Sitti Anggraeni, Darnah A. Nohe...* 334

Pengaruh Tingkat Partisipasi Angkatan Kerja dan Indeks Pembangunan Manusia Terhadap Tingkat Pengangguran Terbuka dengan Regresi Probit dan Logit

*Nabila Abda Salsabila, Sindi Andriani, Mirisda, Darnah Andi Nohe.....* 344

Pemodelan Persentase Penduduk Miskin Di Indonesia Menggunakan Regresi Probit Dan Regresi Logistik

*Grace Wulandari, Nur A. Febriyanti, Khoiril Anwar, Darnah A. Nohe.....* 354

Model Regresi Weibull Pada Data Waktu Rawat Inap Pasien Penderita Penyakit Jantung Koroner Dengan *Event* Kematian di RSUD Abdul Wahab Sjahranie Samarinda

*Hasmiati, Suyitno, Yuki Nova Nasution.....* 369

Analisis Faktor-Faktor Yang Memengaruhi Indeks Pembangunan Manusia Di Jawa Barat Menggunakan Regresi Logistik Biner

*Siti F. Khairunnisa, Suharni, Nurul Fatmi'aturro'isah, Darnah A. Nohe....* 377

Perbandingan Keakuratan Model Logit Dan Probit dalam Mengidentifikasi Faktor-Faktor yang Memengaruhi Persentase Penduduk Miskin di Kalimantan

*Adelia Ramadhani, Dhita Reginna Putri, Risti Raihani, Darnah Andi Nohe...* 387

## PEMILIHAN MODEL REGRESI BERNOULLI TERBAIK

M. Fathurahman<sup>1\*</sup>

<sup>1</sup>Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam,  
Universitas Mulawarman, Indonesia

*Corresponding author:* fathur@fmipa.unmul.ac.id

**Abstrak.** Regresi Bernoulli merupakan model regresi yang dapat menjelaskan hubungan antara variabel dependen kategorik dan satu atau lebih variabel independen. Penelitian ini bertujuan mengkaji pemilihan model regresi Bernoulli terbaik menggunakan metode *Akaike's Information Criterion (AIC)*, *Corrected AIC (AICc)*, *Bayesian Information Criterion (BIC)*. Ketiga metode ini berhubungan dengan metode *Maximum Likelihood Estimation (MLE)* yang digunakan untuk estimasi parameter model regresi Bernoulli. Berdasarkan hasil kajian empiris pada data Indeks Pembangunan Kesehatan Masyarakat (IPKM) kabupaten/kota di Pulau Kalimantan, metode AIC menghasilkan nilai terkecil dibanding metode AICc dan BIC untuk semua model yang dibandingkan. Model regresi Bernoulli terbaik untuk data IPKM berdasarkan nilai AIC, AICc, dan BIC terkecil adalah model dengan variabel independennya adalah produk domestik regional bruto per kapita.

**Kata Kunci:** *regresi Bernoulli, MLE, AIC, AICc, BIC, IPKM.*

## 1 PENDAHULUAN

Regresi Bernoulli merupakan salah satu pendekatan alternatif yang dapat digunakan untuk memodelkan hubungan antara respon kategorik dan satu atau lebih prediktor kategorik, kuantitatif, atau gabungan antara keduanya. Model regresi Bernoulli adalah model marginal dari model regresi Bernoulli bivariat yang dikembangkan oleh [1]. Variabel respon model regresi Bernoulli mempunyai dua kategori dan berdistribusi Bernoulli univariat. Distribusi ini merupakan distribusi marginal dari distribusi Bernoulli bivariat yang ditemukan oleh Marshall dan Olkin pada tahun 1985 dalam [2] dan [3]. Penelitian ini difokuskan untuk model regresi Bernoulli yang mempunyai satu variabel dependen (univariat), khususnya untuk pemilihan model terbaik.

Model regresi terbaik akan memberikan hasil yang optimal dalam estimasi, inferensi, maupun prediksi. Beberapa penelitian yang mengkaji pemilihan model regresi terbaik diantaranya adalah pemilihan model regresi terbaik dan penerapannya pada bidang pendidikan menggunakan metode *Akaike Information Criterion* (AIC) dan *Schwarz Information Criterion* (SIC) [4]. [5] mengkaji dan menerapkan metode AIC dan *Bayesian Information Criterion* (BIC) pada pemilihan model terbaik dengan menggunakan data psikologi. [6] mengembangkan metode *Corrected AIC* (AICc).

Berdasarkan penelitian sebelumnya, penelitian yang mengkaji pemilihan model regresi Bernoulli terbaik menggunakan metode AIC, AICc, dan BIC masih sangat sedikit. Oleh karena itu, penelitian ini bertujuan mengkaji pemilihan model regresi Bernoulli terbaik menggunakan ketiga metode tersebut dan penerapannya pada data Indeks Pembangunan Kesehatan Masyarakat (IPKM) kabupaten/kota di Pulau Kalimantan tahun 2018.

## 2 TINJAUAN PUSTAKA

### 2.1 Distribusi Bernoulli

Distribusi Bernoulli merupakan distribusi khusus dari distribusi binomial [7]. Jika  $Y$  adalah variabel random berdistribusi Bernoulli, maka fungsi masa peluang dari  $Y$  dinyatakan sebagai berikut:

$$P(Y = y) = \gamma^y(1 - \gamma)^{1-y}, \quad y = 0,1. \quad (1)$$

Rata-rata dan varians dari  $Y$  berdasarkan Persamaan (1) adalah berturut-turut  $\mu = E(Y) = \gamma$  dan  $\sigma^2 = Var(Y) = \gamma(1 - \gamma)$ , dimana  $\gamma = P(Y = 1)$  disebut peluang kejadian sukses untuk  $Y$  dan  $\gamma(1 - \gamma) = P(Y = 0)$  disebut peluang kejadian gagal untuk  $Y$ .

### 2.2 Regresi Bernoulli

Regresi Bernoulli merupakan model marginal dari model regresi Bernoulli bivariat [1-3]. Jika diberikan sampel random  $Y_1, Y_2, \dots, Y_n$  berdistribusi Bernoulli dengan parameter  $\gamma$ , maka model regresi Bernoulli dapat ditulis sebagai berikut [1-3]:

$$h(\mathbf{x}_i) = \ln \left[ \frac{\gamma(\mathbf{x}_i)}{1 - \gamma(\mathbf{x}_i)} \right] = \boldsymbol{\theta}^T \mathbf{x}_i, \quad (2)$$

dimana  $\mathbf{x}_i = [1 \quad X_{1i} \quad X_{2i} \quad \dots \quad X_{ki}]^T$  adalah vektor variabel independen untuk pengamatan ke-  $i$ ,  $i = 1, 2, \dots, n$ .  $\boldsymbol{\theta}^T = [\theta_0 \quad \theta_1 \quad \theta_2 \quad \dots \quad \theta_p]$  adalah vektor

parameter.  $\gamma(\mathbf{x}_i)$  adalah peluang variabel dependen yang mempunyai kategori bernilai 1 untuk pengamatan ke- $i$  dan bergantung pada variabel independen pada pengamatan ke- $i$ , yaitu:

$$\gamma(\mathbf{x}_i) = P(Y_i = 1|\mathbf{x}_i) = \frac{\exp(\boldsymbol{\theta}^T \mathbf{x}_i)}{1 + \exp(\boldsymbol{\theta}^T \mathbf{x}_i)}. \quad (3)$$

Untuk mendapatkan model regresi Bernoulli, dilakukan estimasi terhadap parameter modelnya menggunakan metode MLE [1-3]. Estimasi parameter diawali dengan mengambil  $n$  sampel random yang saling independen untuk variabel random  $Y_i$  berdistribusi Bernoulli dengan parameter  $\gamma(\mathbf{x}_i)$ . Berdasarkan Persamaan (1), diperoleh fungsi likelihood sebagai berikut:

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}) &= \prod_{i=1}^n P(Y_i = y_i) \\ &= \prod_{i=1}^n \gamma(\mathbf{x}_i)^{y_i} [1 - \gamma(\mathbf{x}_i)]^{1-y_i} \\ &= \prod_{i=1}^n [\exp(\boldsymbol{\theta}^T \mathbf{x}_i)]^{y_i} [1 + \exp(\boldsymbol{\theta}^T \mathbf{x}_i)]^{-1}. \end{aligned} \quad (4)$$

Selanjutnya, dilakukan transformasi terhadap Persamaan (4) menggunakan fungsi *natural logarithm*, sehingga didapatkan fungsi log-likelihood

$$\ell(\boldsymbol{\theta}) = \ln[\mathcal{L}(\boldsymbol{\theta})] = \sum_{i=1}^n y_i \boldsymbol{\theta}^T \mathbf{x}_i - \sum_{i=1}^n \ln[1 + \exp(\boldsymbol{\theta}^T \mathbf{x}_i)]. \quad (5)$$

Estimator untuk parameter model regresi Bernoulli dapat diperoleh dengan memaksimalkan fungsi log-likelihood pada Persamaan (5) dengan cara menentukan derivatif parsial pertama fungsi log-likelihood terhadap  $\boldsymbol{\theta}$  kemudian disamakan dengan nol, yaitu:

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T} = \sum_{i=1}^n y_i \mathbf{x}_i - \sum_{i=1}^n \mathbf{x}_i \gamma(\mathbf{x}_i) = 0, \quad (6)$$

dimana  $\gamma(\mathbf{x}_i)$  seperti pada Persamaan (3).

Berdasarkan Persamaan (6), diperoleh sistem persamaan yang implisit. Hal ini menunjukkan bahwa estimator parameter model regresi Bernoulli yang diperoleh dengan metode MLE tidak menghasilkan penyelesaian secara analitik. Oleh karena itu, untuk mendapatkan estimator parameter digunakan pendekatan secara numerik menggunakan metode *Fisher scoring* [8,9] dengan formula sebagai berikut:

$$\hat{\boldsymbol{\theta}}^{(t+1)} = \hat{\boldsymbol{\theta}}^{(t)} + \mathbf{I}^{-1}(\hat{\boldsymbol{\theta}}^{(t)}) \mathbf{g}(\hat{\boldsymbol{\theta}}^{(t)}), t = 0, 1, 2 \dots \quad (7)$$

dimana  $\hat{\boldsymbol{\theta}}$  adalah estimator parameter;  $\mathbf{I}^{-1}(\hat{\boldsymbol{\theta}})$  adalah invers dari matriks informasi Fisher, dengan elemen-elemen dari matriks informasi Fisher adalah derivatif parsial kedua fungsi log-likelihood terhadap parameter yang diestimasi; dan  $\mathbf{g}(\hat{\boldsymbol{\theta}})$  adalah vektor gradien dengan elemen-elemennya adalah derivatif parsial pertama fungsi log-likelihood terhadap parameter yang diestimasi, seperti pada Persamaan (6).

Vektor gradien dan matriks informasi Fisher berturut-turut dinyatakan sebagai berikut [8]:

$$\mathbf{g}(\hat{\boldsymbol{\theta}}) = \left[ \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T} \right]^T, \quad (8)$$

$$\mathbf{I}^{-1}(\hat{\boldsymbol{\theta}}^{(t)}) = -E \left[ \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right], \quad (9)$$

dengan  $E \left[ \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right]$  adalah harga harapan dari derivatif parsial kedua fungsi log-likelihood terhadap parameter yang ditaksir.

Berdasarkan Persamaan (7), estimator parameter model regresi Bernoulli dapat diperoleh pada saat proses iterasi *Fisher scoring* memenuhi kondisi konvergen, yaitu  $\|\hat{\boldsymbol{\theta}}^{(t+1)} - \hat{\boldsymbol{\theta}}^{(t)}\| \leq \delta$ , dengan  $\delta$  adalah bilangan positif yang sangat kecil [8]. Estimator parameter yang didapat adalah estimator pada saat iterasi terakhir, yaitu  $\hat{\boldsymbol{\theta}}^{(t+1)}$ .

### 2.3 Pemilihan Model Terbaik

Pemilihan model regresi Bernoulli terbaik dapat dilakukan dengan menggunakan metode AIC, AICc, dan BIC. Metode AIC ditemukan oleh Akaike pada tahun 1974 dan metode BIC ditemukan oleh Schwarz pada tahun 1978 [5]. Sedangkan, metode AICc ditemukan oleh Hurvich dan Tsai pada tahun 1989 [6]. Formula yang digunakan untuk menghitung nilai AIC, AICc, dan BIC adalah sebagai berikut [5,6,10]:

$$\begin{aligned} \text{AIC} &= -2\ell(\hat{\boldsymbol{\theta}}) + 2(p + 1), \\ \text{AICc} &= -2\ell(\hat{\boldsymbol{\theta}}) + \frac{2n(p + 1)}{n - p - 2}, \\ \text{BIC} &= -2\ell(\hat{\boldsymbol{\theta}}) + \ln(n)(p + 1), \end{aligned} \quad (10)$$

dimana  $\ell(\hat{\boldsymbol{\theta}})$  adalah nilai fungsi log-likelihood untuk estimator parameter,  $p$  adalah banyaknya parameter yang diestimasi, dan  $n$  adalah banyaknya sampel. Berdasarkan Persamaan (10), model regresi Bernoulli terbaik adalah model yang mempunyai nilai AIC, AICc, dan BIC terkecil.

## 3 DATA

Data yang digunakan pada penelitian ini adalah data sekunder yang diperoleh dari Badan Penelitian dan Pengembangan Kesehatan (Balitbangkes) Kementerian Kesehatan dan Badan Pusat Statistik. Data yang diambil dari Balitbangkes Kementerian Kesehatan adalah data hasil publikasi IPKM kabupaten/kota tahun 2018. Data yang diambil dari Badan Pusat Statistik adalah data hasil publikasi dan hasil survei. Data hasil publikasi meliputi data dan informasi kemiskinan tahun 2018, provinsi dalam angka, dan laporan pembangunan manusia tahun 2018, sedangkan data hasil survei meliputi hasil Survei Sosial Ekonomi Nasional (Susenas) tahun 2018.

Variabel yang digunakan dalam penelitian ini meliputi variabel dependen dan variabel independen. Variabel dependen ( $Y$ ) adalah status IPKM kabupaten/kota di Pulau Kalimantan tahun 2018. Sedangkan variabel independen ( $X_j$ ) adalah faktor-faktor yang diduga berpengaruh terhadap status IPM kabupaten/kota di Pulau Kalimantan tahun 2018. Variabel independen yang

digunakan dalam penelitian ini mengacu pada penelitian-penelitian sebelumnya mengenai beberapa faktor yang berpengaruh signifikan terhadap status IPKM [8,11,12]. Variabel-variabel yang digunakan dalam penelitian ini disajikan pada Tabel 1.

Tabel 1: Variabel Penelitian

Variabel	Nama Variabel	Skala Pengukuran
$Y$	Status IPKM, dengan kategori [13]: 0 = IPKM DBK ( $IPKM < \text{rata-rata}$ ) 1 = IPKM B-DBK ( $IPKM \geq \text{rata-rata}$ ) DBK = Daerah Bermasalah Kesehatan B-DBK = Bukan Daerah Bermasalah Kesehatan.	Nominal
$X_1$	Status IPM, dengan kategori [14]: 0 = IPM sedang ( $60 \leq IPM < 70$ ) 1 = IPM tinggi ( $70 \leq IPM < 80$ )	Nominal
$X_2$	Persentase penduduk miskin	Rasio
$X_3$	PDRB per kapita	Rasio
$X_4$	Persentase penduduk yang berpendidikan minimal SMP	Rasio

Tahapan-tahapan analisis data dalam penelitian ini adalah sebagai berikut:

- 1) Melakukan analisis statistika deskriptif terhadap variabel dependen dan variabel independen.
- 2) Melakukan pendeteksian multikolinieritas pada variabel independen menggunakan nilai VIF.
- 3) Melakukan estimasi parameter model regresi Bernoulli menggunakan metode MLE dan *Fisher scoring*.
- 4) Melakukan pemilihan model terbaik.
- 5) Melakukan interpretasi terhadap model terbaik.
- 6) Menarik kesimpulan.

#### 4 HASIL DAN PEMBAHASAN

Pembahasan pada penelitian ini diawali dengan analisis statistik deskriptif terhadap variabel penelitian. Deskripsi dari variabel dependen disajikan pada Tabel 2. Sedangkan deskripsi dari variabel independen disajikan pada Tabel 3 dan Tabel 4.

Tabel 2. Deskripsi Variabel Dependen

Status IPKM	Frekuensi	Persentase
IPKM DBK	30	53,57
IPKM B-DBK	26	46,43
Jumlah	56	100

Berdasarkan Tabel 2, terlihat bahwa status IPKM ( $Y$ ) kabupaten/kota di Pulau Kalimantan tahun 2018 terbanyak adalah IPKM DBK, yaitu sebanyak 30 kabupaten/kota (53,57%).



**Tabel 3.** Deskripsi Variabel Independen Kategorik

Status IPM	Frekuensi	Persentase
IPM Sedang	33	58,93
IPM Tinggi	23	41,07
Jumlah	56	100

Pada Tabel 3, tampak bahwa status IPM ( $X_1$ ) kabupaten/kota di Pulau Kalimantan tahun 2018 terbanyak adalah IPM sedang, yaitu sebanyak 33 kabupaten/kota (58,93%).

**Tabel 4.** Deskriptif Variabel Independen Kuantitatif

Variabel	Min	Max	Mean	SD
$X_2$	2,64	12,83	6,27	2,43
$X_3$	23,13	353,30	67,20	76,86
$X_4$	35,58	81,37	54,25	11,05

Berdasarkan Tabel 4, dapat diketahui bahwa rata-rata persentase penduduk miskin ( $X_2$ ) di Pulau Kalimantan tahun 2018 sebesar 6,27%. Persentase penduduk miskin terendah di Kota Balikpapan, Kalimantan Timur sebesar 2,64% dan persentase tertinggi di Kabupaten Melawi, Kalimantan Barat sebesar 12,83%. Rata-rata PDRB per kapita ( $X_3$ ) kabupaten/kota di Pulau Kalimantan tahun 2018 sebesar 67,2 juta rupiah, dengan PDRB per kapita terendah di Kabupaten Melawi, Kalimantan Barat sebesar 23,13 juta rupiah dan PDRB per kapita tertinggi di Kabupaten Kutai Timur, Kalimantan Timur sebesar 353,3 juta rupiah. Rata-rata persentase penduduk yang berpendidikan minimal SMP ( $X_4$ ) di Pulau Kalimantan tahun 2018 sebesar 54,25%. Persentase terendah di Kabupaten Sambas, Kalimantan Barat sebesar 35,58%. Sedangkan, persentase tertinggi di Kota Palangkaraya, Kalimantan Tengah sebesar 81,37%.

Pembahasan selanjutnya adalah melakukan pendeteksian multikolinieritas terhadap variabel independen sebagai prasyarat analisis data pada model regresi Bernoulli. Pendekatan yang digunakan untuk mendeteksi adanya multikolinieritas adalah *Variance Inflation Factor* (VIF). Model regresi Bernoulli bebas masalah multikolinieritas jika semua variabel independennya mempunyai nilai VIF kurang dari 10 [15]. Berdasarkan hasil perhitungan dengan *software* R, diperoleh nilai VIF untuk setiap variabel independen yang disajikan pada Tabel 5.

**Tabel 5.** Nilai VIF Variabel Independen

Variabel Independen	Nilai VIF
$X_2$	1,25
$X_3$	1,37
$X_4$	1,62

Pada Tabel 5, tampak bahwa variabel  $X_2$ ,  $X_3$ , dan  $X_4$  mempunyai nilai VIF kurang dari 10. Hasil ini menunjukkan bahwa tidak terdapat multikolinieritas. Oleh karena itu, variabel independen  $X_2$ ,  $X_3$ , dan  $X_4$  dapat digunakan untuk model regresi Bernoulli.

Pemilihan model regresi Bernoulli terbaik untuk data IPKM kabupaten/kota di Pulau Kalimantan tahun 2018, diawali dengan melakukan estimasi terhadap

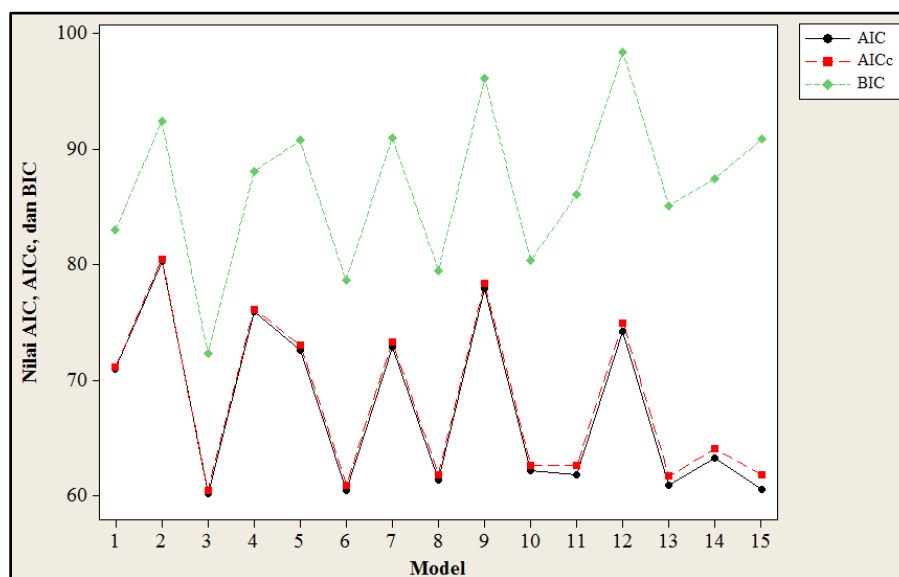
parameter model regresi Bernoulli menggunakan metode MLE dan iterasi *Fisher scoring* seperti pada Persamaan (7). Selanjutnya, menghitung nilai AIC, AICc, dan BIC berdasarkan Persamaan (10). Setelah dilakukan perhitungan dengan *software R*, diperoleh hasil yang disajikan pada Tabel 6.

**Tabel 6.** Nilai AIC, AICc, dan BIC Model Regresi Bernoulli Data IPKM

Model	Variabel Independen	AIC	AICc	BIC
1	$X_1$	70,94	71,17	83,04
2	$X_2$	80,30	80,52	92,40
3	$X_3$	60,23*	60,45*	72,33*
4	$X_4$	75,96	76,18	88,06
5	$X_1X_2$	72,62	73,08	90,77
6	$X_1X_3$	60,51	60,97	78,66
7	$X_1X_4$	72,86	73,33	91,02
8	$X_2X_3$	61,37	61,84	79,53
9	$X_2X_4$	77,96	78,42	96,11
10	$X_3X_4$	62,22	62,69	80,38
11	$X_1X_2X_3$	61,87	62,65	86,07
12	$X_1X_2X_4$	74,22	75,00	98,42
13	$X_1X_3X_4$	60,91	61,70	85,12
14	$X_2X_3X_4$	63,27	64,06	87,48
15	$X_1X_2X_3X_4$	60,59	61,79	90,85

\*) Model terbaik

Berdasarkan Tabel 6, terlihat bahwa model regresi Bernoulli dengan variabel independen PDRB per kapita (Model 3) adalah model terbaik untuk data IPKM. Hal ini ditunjukkan oleh nilai AIC, AICc, dan BIC untuk Model 3 terkecil dibanding model-model lainnya. Hasil ini juga disajikan pada Gambar 1.



**Gambar 1.** Nilai AIC, AICc, dan BIC Model Regresi Bernoulli Data IPKM

Model regresi Bernoulli terbaik yang diperoleh dari Tabel 5 dapat ditulis sebagai berikut:

$$\hat{h}(x) = -1,82 + 0,03X_3. \quad (11)$$

Interpretasi dari model pada Persamaan (11) adalah setiap penambahan PDRB per kapita sebesar 1 juta rupiah, maka kabupaten/kota memiliki kecenderungan sebesar 1,03 kali mempunyai status IPKM B-DBK daripada IPKM DBK.

## 5 KESIMPULAN

Pemilihan model regresi Bernoulli terbaik dapat dilakukan dengan menggunakan metode AIC, AICc, dan BIC karena ketiga metode ini bergantung pada fungsi log-likelihood dari metode MLE yang digunakan untuk estimasi parameter model regresi Bernoulli. Performa dari metode AIC, AICc, dan BIC dievaluasi dengan menerapkannya pada data IPKM kabupaten/kota di Pulau Kalimantan tahun 2018. Berdasarkan hasil penerapan pada data IPKM, metode AIC memberikan nilai terkecil dibanding metode AICc dan BIC. Model regresi Bernoulli terbaik untuk data IPKM kabupaten/kota di Pulau Kalimantan tahun 2018 adalah model regresi Bernoulli dengan variabel independennya adalah PDRB per kapita.

## DAFTAR PUSTAKA

- [1] Bhuyan, M. B., Islam, M. A., & Rahman, M. S. (2018). A Bivariate Bernoulli Model for Analyzing Malnutrition Data. *Health Services and Outcomes Research Methodology*, 18, 109-127.
- [2] Islam, M. A., Chowdhury, R. I., & Briollais, L. (2012). A Bivariate Binary Model for Testing Dependence in Outcomes. *Bulletin of the Malaysian Mathematical Sciences Society*, 35(4), 845-858.
- [3] El-Sayed, A. M., Islam, M. A., & Alzaid, A. A. (2013). Estimation and Test of Measures of Association for Correlated Binary Data. *Bulletin of the Malaysian Mathematical Sciences Society*, 36(4), 985-1008.
- [4] Fathurahman, M. (2009). Pemilihan Model Regresi Terbaik Menggunakan Metode Akaike's Information Criterion dan Schwarz Information Criterion. *Jurnal Informatika Mulawarman*, 4(3), 37-41.
- [5] Vrieze, S. I. (2012). Model Selection and Psychological Theory: A discussion of the differences between the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). *Psychol Methods*, 17(2), 228-243.
- [6] Delsole, T., & Tippett, M. K. (2021). Correcting the Corrected AIC. *Statistics and Probability Letters*, 173, 109064.
- [7] Johnson, N. L., Kemp, A. W., & Kotz, S. (2005). *Univariate Discrete Distributions* (3rd ed.). New Jersey: John Wiley & Sons.
- [8] Fathurahman, M. (2019). Pemodelan Indeks Pembangunan Kesehatan Masyarakat Kabupaten/Kota di Pulau Kalimantan Menggunakan Pendekatan Regresi Probit. *Jurnal Varian*, 2(2), 47-54.
- [9] Fathurahman, M., Siringoringo, M., Satriya, A. M., & Sari, N. W. (2019). Pemodelan Regresi Logistik dan Regresi Probit pada Indeks Pembangunan Manusia Kabupaten/Kota di Pulau Kalimantan. *Prosiding SNMSA* (pp. 172-182). Samarinda: Jurusan Matematika FMIPA Universitas Mulawarman.
- [10] Bilder, C. R., & Loughin, T. M. (2015). *Analysis of Categorical Data with R*. Boca Raton: CRC Press.

- [11] Fibriyani, V., Latra, I. N., & Puhadi. (2015). Geographically Weighted Multinomial Logistic Regression Model (Case Study: Human Development Index Value and Healths Status Areas of Districts/Cities 2013 in Sumatera). *Proceedings of the IConSSE* (pp. 13-19). Salatiga, Indonesia: FSM SWCU.
- [12] Fathurahman, M., Hayati, M. N., & Rizki, N. A. (2019). Pemodelan Indeks Pembangunan Kesehatan Masyarakat Kabupaten/Kota di Pulau Kalimantan dengan Regresi Spasial. *Prosiding SNMSA* (pp. 183-191). Samarinda: Jurusan Matematika FMIPA Universitas Mulawarman.
- [13] Kementerian Kesehatan. (2010). *Pedoman Umum Penanggulangan Daerah Bermasalah Kesehatan Kabupaten/Kota*. Jakarta: Kementerian Kesehatan.
- [14] Badan Pusat Statistik. (2019). *Indeks Pembangunan Manusia 2018*. Jakarta: BPS.
- [15] Kutner, M. H., Nachtsheim, C. J., & Neter, J. (2004). *Applied Linear Regression Models*. New York: McGraw-Hill/Irwin.

Jurusan Matematika  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Universitas Mulawarman  
Jalan Barong Tongkok No 4 Kampus Gunung Kelua  
Samarinda 75123 Kalimantan Timur

