

## **Analisis Biplot Terhadap Aspek Penilaian Kesehatan pada Bank Konvensional**

Selvi Mayang Sari<sup>1</sup>, Sigit Nugroho<sup>2</sup>, dan Idhia Sriliana<sup>2</sup>

<sup>1</sup>Alumni Jurusan Statistika FMIPA Universitas Bengkulu

<sup>2</sup>Staf Pengajar Jurusan Statistika FMIPA Universitas Bengkulu

Abstrak-Penelitian ini bertujuan untuk mengaplikasikan analisis biplot terhadap aspek penilaian kesehatan pada bank konvensional. Data yang digunakan pada penelitian ini berupa data sekunder yang diperoleh dari data laporan posisi keuangan yang diterbitkan oleh Bank Indonesia. Hasil penelitian ini menunjukkan bahwa bank BTN dan BRI diposisikan sebagai objek terlemah untuk semua variabel yang ada. Sedangkan bank yang mempunyai posisi terbaik adalah bank BNI dan Mandiri.

Kata Kunci : *Analisis Biplot, Bank Konvensional.*

### **I. PENDAHULUAN**

Bank merupakan perusahaan yang bergerak dalam bidang keuangan, artinya aktivitas perbankan selalu berkaitan dalam bidang keuangan. Sehingga berbicara mengenai bank tidak terlepas dari masalah keuangan (Kasmir, 2008). Sebagai lembaga perantara keuangan, bank memiliki fungsi menghimpun dana dari pihak yang memiliki kelebihan dana dan menyalurkannya kepada pihak yang memerlukan dana (Neni, 2008).

Perkembangan bank juga tidak hanya terdapat pada kualitas fungsi dan peranannya saja, tetapi juga pada kuantitas jumlah bank itu sendiri, semakin banyaknya jenis dan macam bank konvensional ini sudah tentu persaingan di industri perbankan akan semakin ketat. Semakin banyaknya bank-bank swasta yang beroperasi menyebabkan bank pemerintah harus bekerja keras untuk mempertahankan eksistensinya. Dalam hal ini bank perlu dinilai kesehatannya karena kegiatan bank berhubungan dengan dana-dana yang berasal dari masyarakat dan kegiatan usahanya sesuai dengan prinsip kepercayaan dari nasabahnya (Pandia, 2012). Menurut Mattjik dan Sumertajaya (2011), dalam ilmu statistika, banyak metode yang dapat digunakan untuk mendeskripsikan posisi relatif

beberapa objek dengan beberapa variabel atau peubah secara serempak. Salah satu metodenya adalah dengan menggunakan Analisis Biplot. Biplot adalah salah satu upaya menggambarkan data-data yang ada pada tabel ringkasan dalam grafik berdimensi dua, informasi yang diberikan oleh biplot mencakup objek dan peubah dalam satu gambar (Johnson dan Wichern, 2002).

Tujuan dari penelitian ini adalah pengaplikasian analisis biplot terhadap aspek penilaian kesehatan pada bank konvensional.

### **II. TINJAUAN PUSTAKA**

#### **2.1 Bank**

Bank perlu dinilai kesehatannya karena kegiatan bank berhubungan dengan dana-dana yang berasal dari masyarakat dan kegiatan usahanya sesuai dengan prinsip kepercayaan dari nasabahnya. Tata cara penilaian kesehatan bank yang baru menurut peraturan No. 6/10/PBI/2004 dengan pendekatan CAMELS, berikut ini akan dijelaskan secara rinci untuk masing-masing komponen, yaitu sebagai berikut (Rivai, 2007):

1. Penilaian Permodalan (Capital-C), Capital, untuk memastikan kecukupan modal dan cadangan untuk memikul resiko yang mungkin timbul

2. Kualitas aktiva produktif (KAP) yaitu perbandingan antara *classified assets* (kredit kurang lancar, kredit diragukan dan kredit macet) dengan total *earning assets* (kredit yang diberikan, surat berharga, aktiva antarbank dan penyertaan).
3. Penilaian Kualitas Manajemen Manajemen untuk memastikan kualitas dan tingkat kedalaman penerapan prinsip manajemen bank yang sehat, terutama yang terkait dengan manajemen umum dan manajemen risiko.
4. Penilaian Rentabilitas (*Earning*) ROA merupakan rasio perbandingan antara laba sebelum pajak dengan total aset. Rasio ini digunakan untuk mengukur kemampuan bank dalam memperoleh keuntungan secara keseluruhan, Selain ROA, penilaian dapat dilakukan melalui rasio beban operasional terhadap pendapatan operasional (BOPO). Rasio ini merupakan perbandingan antara biaya operasional dengan pendapatan operasional dalam mengukur tingkat efisiensi dan kemampuan bank dalam melakukan kegiatan operasinya.
5. Penilaian Likuiditas (*Liquidity*) LDR merupakan rasio yang mengukur perbandingan jumlah kredit yang diberikan bank dengan dana yang diterima oleh bank, yang menggambarkan kemampuan bank dalam membayar kembali penarikan dana oleh deposan dengan mengandalkan kredit yang diberikan sebagai sumber likuiditasnya 1.

## 2.2 Matriks

Secara umum, sebuah matriks  $A$  yang berukuran  $m$  baris dan  $n$  kolom dengan elemen  $a_{ij}$  dapat dituliskan sebagai berikut:

$$A_{m \times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

di mana  $i = 1, 2, \dots, m$  dan  $j = 1, 2, \dots, n$ .

## 2.3 Analisis Komponen Utama

Di dalam statistika terdapat usaha-usaha untuk menyederhanakan struktur data atau mereduksi dimensi data tanpa mengabaikan variabel-variabel yang telah diukur tersebut. Hal ini dilakukan dengan jalan menghilangkan korelasi diantara peubah melalui transformasi peubah asal ke peubah baru (merupakan kombinasi linier dari peubah-peubah asal) yang tidak saling berkorelasi. dan  $p$  peubah asal dapat dibentuk  $p$  komponen utama, dipilih  $k$  komponen utama saja ( $k < p$ ) yang telah mampu menerangkan keragaman data cukup tinggi (antara 80% sampai dengan 90%). Komponen yang dipilih tersebut telah dapat mengganti  $p$  peubah asal tanpa mengurangi informasi (Widiharih, 2001).

## 2.4 Analisis Biplot

Biplot adalah salah satu upaya menggambarkan data-data yang ada pada tabel ringkasan dalam grafik berdimensi dua. Informasi yang diberikan oleh biplot mencakup objek dan peubah dalam satu gambar. Empat hal penting yang bisa didapatkan dari tampilan biplot sebagai berikut (Mattjik dan Sumertajaya, 2011).

1. Kedekatan Antar Objek yang Diamati, Kedekatan antar objek diinterpretasikan sebagai panduan untuk mengetahui objek yang memiliki kemiripan karakteristik dengan objek lain, Kedekatan antar objek pada gambar biplot dapat dilihat dengan menggunakan jarak Euclid antara

$g_i$  dan  $g_j$ . Jarak Euclid antara dua pengamatan  $g_i$  dan  $g_j$  didefinisikan sebagai berikut :

$$d^2(g_i, g_j) = (g_i - g_j)'(g_i - g_j)$$

2. Keragaman Peubah, Keragaman peubah digunakan untuk melihat apakah ada variabel yang mempunyai nilai keragaman yang hampir sama untuk setiap objek Dalam biplot, variabel yang mempunyai nilai keragaman yang kecil digambarkan sebagai vektor pendek sedangkan variabel dengan nilai keragaman yang besar digambarkan sebagai vektor yang panjang
3. Korelasi antar peubah digunakan untuk mengetahui bagaimana suatu variabel mempengaruhi ataupun dipengaruhi variabel yang lain. Pada biplot, variabel akan digambarkan sebagai garis berarah.
4. Nilai Peubah pada suatu objek Informasi ini digunakan untuk melihat keunggulan dari setiap objek. Objek yang terletak searah dengan arah vektor variabel dikatakan bahwa objek tersebut mempunyai nilai di atas rata-rata. Namun jika objek terletak berlawanan dengan arah dari vektor variabel tersebut, maka objek tersebut memiliki nilai di bawah rata-rata.

#### 2.4.1. Penguraian Nilai SVD/ Singular Value Decomposition

Menurut Jolliffe (2010), SVD bertujuan menguraikan suatu matriks  $X$  berukuran  $n \times p$  yang merupakan matriks peubah ganda yang terkoreksi terhadap rataannya dimana  $n$  adalah banyaknya objek pengamatan dan  $p$  adalah banyaknya peubah, sehingga analisis model ini dapat disajikan dalam grafik berdimensi 2 atau disebut dengan biplot. diperoleh :

$$X_{n \times p} = {}_n U_r L_r A'_p$$

Dimana

$X_{n \times p}$ : Matriks yang terkoreksi dengan nilai tengah.

$U_{n \times r}$  : Matriks vektor eigen dari  $XX'$

$L_{r \times r}$  : Matriks diagonal akar dari nilai eigen  $XX'$

$A'_{r \times p}$  : Matriks eigen dari  $X'X$

Menurut Jolliffe (2010), misalkan  $G = UL^\alpha$  dan  $H' = L^{1-\alpha}A'$  dengan  $\alpha$  besarnya  $0 \leq \alpha \leq 1$ .

$$X = UL^\alpha L^{1-\alpha} A' = GH'$$

#### 2.4.2. Transformasi Data

Dalam Biplot XLS terdapat beberapa definisi awal untuk transformasi yang mempunyai tipe data yang berbeda-beda (Nugroho, 2008)

- a. Analisis Komponen Utama (AKU) tanpa normalisasi. Transformasi ini dimulai dengan pemusatan kolom-kolom SVD, dan pembuatan Biplot
- b. Analisis Komponen Utama (AKU) dengan normalisasi. Transformasi ini dimulai dengan pemusatan kolom-kolom Normalisasi dari kolom-kolom SVD, dan pembuatan Biplot.
- c. Analisis Koresponden Sederhana. Transformasi ini diawali dengan melakukan transformasi rasio kontingensi Pearson penerapan pembobot baris dan kolom pemusatan ganda terboboti SVD yang terboboti, dan pembuatan Biplot.
- d. *Logratio analysis*. Transformasi ini diawali dengan melakukan transformasi logaritma pemusatan ganda SVD, dan pembuatan Biplot

- e. *Ratio maps*. Transformasi ini diawali dengan melakukan transformasi logaritma penerapan pembobot baris dan kolom pemusatan ganda terboboti SVD terboboti, dan pembuatan Biplot.

#### 2.4.2. Ukuran Ketepatan Biplot

Menurut Mattjik dan Sumertajaya (2011), ukuran ketepatan biplot pada matriks  $X$  dalam bentuk:

$$\rho^2 = \frac{(\lambda_1 + \lambda_2)}{\sum_{k=1}^m \lambda_k} \times 100\%$$

Dengan  $\rho^2$  adalah ukuran ketepatan biplot dua dimensi,  $\lambda_1$  adalah nilai eigen terbesar pertama,  $\lambda_2$  adalah nilai eigen terbesar kedua, dan  $\lambda_k$  adalah nilai eigen terbesar ke- $k$  dengan  $k = 1, 2, \dots, r$ . apabila  $\rho^2 \geq 70\%$  maka biplot sudah memberikan penyajian yang baik mengenai informasi data yang sebenarnya.

### III. METODE PENELITIAN

#### 3.1 Jenis Penelitian

Penelitian ini merupakan penelitian statistik terapan, yaitu suatu penelitian yang dilakukan dengan mengaplikasikan metode statistika ke dalam bidang-bidang tertentu. Pada penelitian ini, analisis biplot digunakan untuk menganalisis aspek penilaian kesehatan pada bank konvensional

#### 3.2 Pengambilan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder. Penelitian ini menggunakan data Laporan Posisi Keuangan yang diterbitkan oleh Bank Indonesia yang berasal dari website [www.bi.go.id](http://www.bi.go.id). Variabel yang digunakan yaitu: CAR, ROA, BOPO, LDR.

#### 3.3 Variabel Penelitian

Adapun beberapa variabel yang digunakan untuk penelitian yaitu:

1. Variabel Dependen : jarak antara objek penelitian
2. Variabel Karakteristik

Nilai *Capital Adequacy Ratio* (CAR), nilai *Return on Asset* (ROA), nilai Biaya Operasional terhadap Pendapatan Operasional (BOPO), nilai *Loan to Deposit Ratio* (LDR)

(2.33)

#### 3.4 Teknik Analisis Data

Tahapan-tahapan yang akan dilakukan dalam penelitian ini adalah :

1. Pendeskripsian data
2. Menentukan nilai matriks  $X^*$  pada variabel karakteristik
3. Transformasi matriks  $X^*$
4. Analisis Komponen Utama
5. Menentukan matriks *Singular Value Decomposition* (SVD)
6. Menghitung matriks  $G = UL^\alpha$  serta  $H' = L^{1-\alpha}A'$
7. Menelusuri ketepatan biplot
8. Interpretasi Output
9. Kesimpulan

### IV. HASIL DAN PEMBAHASAN

#### 4.1 Deskripsi Data Aspek Penilaian Kesehatan pada Bank Konvensional

Data pada penelitian ini merupakan data sekunder yang diperoleh data laporan posisi keuangan yang diterbitkan oleh bank Indonesia pada Lampiran (1). Pada penelitian ini terdapat beberapa variabel peubah karakteristik yaitu LDR, BOPO, ROA, CAR. Berikut penjelasan secara deskriptif mengenai masing-masing variabel

Tabel 5. Data Aspek Penilaian Kesehatan Bank Konvensional

Objek	LDR	BOPO	ROA	CAR
<b>BNI</b>	77.5	70.99	2.67	16.6
<b>BTN</b>	79.8	59.92	4.41	16.9
<b>BRI</b>	78.0	69.93	3.20	15.4
<b>Mandir</b>	99.5	80.73	1.66	18.0

Berdasarkan perhitungan nilai variabel pada tabel 5, bank yang mampu menjaga LDR berada dibawah 98% yaitu bank BNI, BTN dan BRI sedangkan bank Mandiri mampu menjaga LDR berada di atas 98% sehingga berdasarkan kriteria penilaian LDR, bank BNI, BTN dan BRI dikategorikan sangat sehat dan bank Mandiri termasuk dalam kategori sehat artinya semakin rendah kemampuan likuiditas bank yang bersangkutan maka kemungkinan suatu bank dalam kondisi bermasalah akan semakin kecil. Kemudian bank yang mampu menjaga BOPO berada dibawah 83% yaitu pada bank BNI, BTN, BRI dan Mandiri sehingga berdasarkan kriteria penilaian BOPO, keempat bank tersebut dikategorikan sangat sehat hal ini menunjukkan rasio BOPO dalam melakukan kegiatan operasionalnya semakin efisien karena biaya yang dikeluarkan lebih kecil dibandingkan pendapatan yang diterima. Selanjutnya bank yang mampu menjaga ROA berada di atas 1.5% yaitu pada bank BNI, BTN, BRI dan Mandiri berdasarkan kriteria penilaian ROA keempat bank tersebut termasuk dalam kategori sangat sehat hal ini menunjukkan ROA mampu dalam mengelola aset bank yang dimiliki untuk menghasilkan laba dengan baik. Selanjutnya bank yang mampu menjaga CAR berada di atas 11% yaitu pada bank BNI, BTN,

BRI dan Mandiri sehingga berdasarkan kriteria penilaian CAR keempat bank tersebut termasuk dalam kategori sangat sehat yang artinya semakin besar rasio CAR yang dimiliki oleh bank maka akan semakin baik hal ini dikarenakan bank mampu menyediakan modal dalam jumlah yang besar.

#### 4.2 Menentukan Nilai Matriks X pada Kariabel Karakteristik

Nilai variabel karakteristik ini diperoleh dari pengolahan data mentah laporan posisi keuangan yang diterbitkan oleh Bank Indonesia menggunakan metode CAMELS yang diperoleh pada Lampiran (2), kemudian data pada Tabel 5 dibentuk menjadi matriks, sebagai berikut :

$$X^* = \begin{bmatrix} 77.52 & 70.99 & 2.67 & 16.67 \\ 79.85 & 59.92 & 4.41 & 16.94 \\ 78.07 & 69.93 & 3.20 & 15.47 \\ 99.50 & 80.73 & 1.66 & 18.04 \end{bmatrix}$$

#### 4.3 Transformasi Matriks X

Transformasi yang digunakan yaitu transformasi Analisis Komponen Utama tanpa normalisasi dengan mengurangi nilai data matriks terhadap nilai rata-ratanya. Hasil transformasi matriks dapat dilihat pada Tabel 7 sebagai berikut

Tabel 7. Transformasi Matriks X\*

Objek	LDR	BOPO	ROA	CAR
<b>BNI</b>	-6.2	0.5975	-0.3	-0.1
<b>BTN</b>	-3.8	-10.47	1.42	0.16
<b>BRI</b>	-5.6	-0.462	0.21	-1.3
<b>Mandir</b>	15.7	10.33	-1.32	1.26

#### 4.4 Analisis Komponen Utama

Analisis Komponen Utama bertujuan mengetahui seberapa besar keragaman data

secara keseluruhan dapat dijelaskan oleh kedua komponen utama sebagai berikut

Tabel 8. Analisis Komponen Utama

Variabel	Komponen	
	1	2
<b>LDR</b>	0.79528221	-0.5958145
<b>BOPO</b>	0.59830801	0.7874639
<b>ROA</b>	-0.07760861	-0.1163523
<b>CAR</b>	0.05941921	-0.1066199

Berdasarkan tabel 8 di atas dua komponen utama merupakan kombinasi linier dari variabel-variabel asal. Persamaan komponen utama yang pertama, yaitu  $Y_1 = 0.79528221 LDR + 0.59830801 BOPO - 0.0776086 ROA + 0.05941921 CAR$

Persamaan komponen utama yang kedua  $Y_2 = -0.5958145 LDR + 0.7874639 BOPO - 0.1163523 ROA - 0.1066199 CAR$

Nilai dari  $Y_1$  merupakan komponen utama pertama yang mempunyai keragaman terbesar pertama terhadap data. Besar nilai keragaman yang dapat dijelaskan dari komponen utama pertama adalah 87.9%. Sedangkan nilai dari  $Y_2$  merupakan komponen utama yang mempunyai keragaman terbesar kedua terhadap data. Besar nilai keragaman yang dapat dijelaskan dari komponen utama kedua adalah 11.9%.

#### 4.5 Singular Value Decomposition

nilai *Singular Value Decomposition* dapat diuraikan sebagai berikut :

$$X_{n \times p} = U_{n \times r} L_{r \times r} A'_{r \times p}$$

Diperoleh matriks

$$U = \begin{bmatrix} -0.206 & 0.5176 & -0.66 & 0.501 \\ -0.426 & -0.749 & -0.07 & 0.500 \\ -0.220 & 0.3832 & 0.746 & 0.497 \\ 0.8526 & -0.151 & -0.00 & 0.500 \end{bmatrix}$$

$$L = \begin{bmatrix} 22.16504 & 0 & 0 & 0 \\ 0 & 8.155091 & 0 & 0 \\ 0 & 0 & 0.966159 & 0 \\ 0 & 0 & 0 & 0.009503 \end{bmatrix}$$

$$A' = \begin{bmatrix} 0.79528 & 0.59830 & -0.0776 & 0.05941 \\ -0.5958 & 0.78746 & -0.1163 & -0.1066 \\ 0.1050 & -0.0086 & 0.27828 & -0.9547 \\ -0.0387 & 0.14782 & 0.9502 & 0.27138 \end{bmatrix}$$

#### 4.6. Menghitung Matriks G dan Matriks H'

Matriks G dan matriks H' diperoleh :

$$G = UL$$

$$G = \begin{bmatrix} -4.56767 & 4.2212 & -0.639 & 0.004 \\ -9.45693 & -6.115 & -0.071 & 0.0047 \\ -4.8769 & 3.1251 & 0.7209 & 0.0047 \\ 18.89994 & -1.233 & -0.004 & 0.0047 \end{bmatrix}$$

$$H' = LA'$$

$$H' = \begin{bmatrix} 17.627 & 13.261 & -1.72 & 1.3170 \\ -4.858 & 6.421 & -0.948 & -0.869 \\ 0.1014 & -0.008 & 0.2688 & -0.922 \\ -0.0003 & 0.0014 & 0.009 & 0.0025 \end{bmatrix}$$

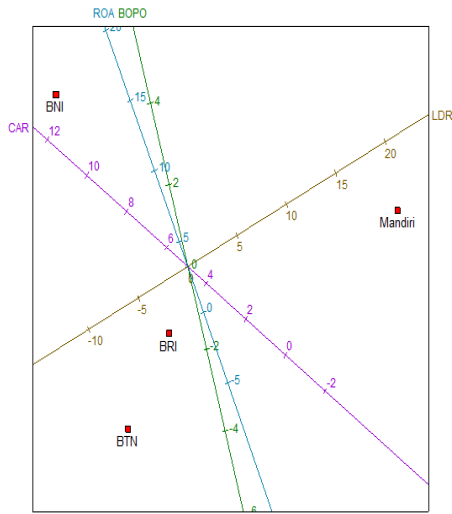
Selanjutnya akan diambil dua kolom pertama dari matriks G dan H' untuk membentuk suatu gambar analisis biplot

$$G = \begin{bmatrix} -4.56767 & 4.221294 \\ -9.45693 & -6.11542 \\ -4.8769 & 3.125161 \\ 18.89994 & -1.23336 \end{bmatrix}$$

$$H' = \begin{bmatrix} 17.62746 & 13.26152 \\ -4.85892 & 6.42184 \\ 0.101466 & -0.00839 \\ -0.00037 & 0.001405 \end{bmatrix}$$

Berdasarkan perhitungan matriks G dan Matriks H' maka didapat gambaran analisis biplot terhadap aspek penilaian kesehatan pada

bank konvensional sehingga diperoleh gambar sebagai berikut



Gambar 2. Analisis Biplot terhadap aspek penilaian kesehatan pada bank konvensional

#### 4.7 Menelusuri Ketepatan Biplot

ketepatan biplot yang bertujuan untuk mengetahui penyajian informasi mengenai data yang sebenarnya. Berdasarkan persamaan (2.33) berikut uraiannya:

$$\rho^2 = \frac{(\lambda_1 + \lambda_2)}{\sum_{k=1}^m \lambda_k} \times 100\%$$

$$= \frac{(22.16504 + 8.155091)}{22.16504 + 8.155091 + 0.966159 + 0.009503} \times 100\% = 96.88\%$$

Nilai  $\lambda_1, \lambda_2$  diperoleh dari matriks diagonal  $L$ , berdasarkan hasil yang diperoleh nilai ketepatan biplot yang dapat diterangkan oleh analisis komponen biplot dua dimensi adalah 96.88% yang berarti Gambar (2) mampu menerangkan 96.88% dari total data yang sebenarnya. Hal ini menunjukkan bahwa interpretasi yang dihasilkan mampu

menerangkan dengan baik hubungan antara bank milik pemerintahan ini dengan variabel aspek penilaian kesehatannya.

#### 4.8 Interpretasi Output

##### 1. Kedekatan antar objek yang diamati

Berdasarkan Gambar 2 diperoleh informasi bahwa kedekatan antar objek ditunjukkan oleh bank BTN dengan bank BRI, yang artinya objek tersebut memiliki kemiripan karakteristik sedangkan objek yang tidak memiliki kedekatan terletak pada objek BNI dengan BTN, BNI dengan BRI, BNI dengan Mandiri, dan BTN dengan Mandiri yang artinya objek tersebut tidak memiliki kemiripan karakteristik.

##### 2. Keragaman Peubah

Keragaman yang dijelaskan pada Gambar 2 menunjukkan bahwa keragaman yang terbesar terjadi pada variabel LDR dan ROA karena memiliki vektor yang paling panjang dibandingkan dengan vektor yang lain. Hal ini menunjukkan bahwa bank-bank pemerintahan tersebut beragam dalam hal penilaian LDR dan ROA. Sedangkan untuk keragaman terkecil terjadi pada variabel BOPO dan CAR karena memiliki vektor yang pendek. Hal ini menunjukkan bank milik pemerintahan tersebut tidak beragam dalam hal penilaian BOPO dan CAR.

##### 3. Korelasi Antar Peubah

Korelasi positif terjadi pada variabel ROA dan BOPO karena kedua variabel tersebut memiliki arah yang sama dan

membentuk sudut lancip yang dimiliki objek pada bank tersebut, yang artinya rasio hubungan antar bank dengan aspek penilaian kesehatan saling berhubungan satu sama lain. Sedangkan untuk variabel LDR dan CAR tidak menunjukkan adanya hubungan satu sama lain karena tidak memiliki arah yang sama.

#### 4. Nilai Peubah pada Suatu Objek

Pada biplot terlihat bahwa bank BTN dan BRI terletak berlawanan dengan arah dari semua vektor vektor variabel yang artinya kedua bank tersebut mempunyai nilai dibawah rata-rata atau dengan kata lain kedua bank tersebut dianggap sebagai bank konvensional yang lemah atau tidak memiliki keunggulan dalam hal aspek penilaian bank. Sedangkan untuk bank Mandiri terletak searah dengan vektor variabel LDR dan bank BNI terletak searah dengan vektor variabel CAR, ROA, dan BOPO yang artinya kedua bank tersebut dianggap sebagai bank konvensional yang memiliki keunggulan dalam hal aspek penilaian bank.

## V. HASIL DAN PEMBAHASAN

### 5.1 Kesimpulan

Berdasarkan dari hasil dan pembahasan yang telah diuraikan pada bab sebelumnya, maka diperoleh informasi sebagai berikut:

1. Analisis Biplot dapat menghasilkan ketepatan sebesar 96.88% dari total data pada bank konvensional.
2. Analisis Biplot memberikan keseluruhan informasi mengenai kedekatan objek yang diamati, keragaman peubah, korelasi antar peubah, dan nilai peubah pada suatu objek
  - a. Objek-objek yang dikategorikan memiliki kedekatan antar objek terletak pada objek BTN dengan BRI, sedangkan objek yang tidak memiliki kedekatan terletak pada objek BNI dengan BTN, BNI dengan BRI, BNI dengan Mandiri, dan BTN dengan Mandiri
  - b. Keragaman yang terbesar terjadi pada variabel LDR dan ROA. Sedangkan untuk keragaman terkecil terjadi pada variabel BOPO dan CAR
  - c. Korelasi positif terjadi pada variabel ROA dan BOPO. Sedangkan untuk variabel LDR dan CAR tidak menunjukkan adanya korelasi positif
  - d. Keunggulan dalam hal aspek penilaian kesehatan bank terletak pada objek BNI dan Mandiri, sedangkan objek BTN dan BRI tidak memiliki keunggulan dalam hal aspek penilaian kesehatan bank.

## 5.2 Saran

Berdasarkan dari hasil penelitian, disarankan untuk melakukan penelitian lanjutan dengan menggunakan metode analisis biplot yang lain.

### DAFTAR PUSTAKA

- (1) A.A Mattijk, dan I.M, Sumertajaya. 2011. *Sidik Peubah Ganda*. IPB Press. Bogor.
- (2) Johnson, R. A, and D. E, Wichern. 2002. *Applied Multivariate Statistical Analysis, Fifth Edition*. Pearson Education International. New York
- (3) Jolliffe, I.T. 2010. *Principal Component Analysis, Second Edition*. Springer. New York .
- (4) Kasmir. 2008. *BANK dan Lembaga Keuangan Lainnya*. PT.RajaGrafindo Persada. Jakarta
- (5) Neni, I.S. 2008. *Hukum Perbankan Untuk Lingkungan Sendiri*. Fakultas Hukum Unisba. Bandung
- (6) Nugroho, S. 2008. *Statistika Multivariat Terapan*. UNIB Press. Universitas Bengkulu
- (7) Pandia, F. 2012. *Manajemen Dana dan Kesehatan Bank*. Rineka Cipta. Jakarta
- (8) Rivai, V. 2007. *Bank and Financial Institution Management*. PT RajaGrafindo Persada. Jakarta
- (9) Widiharih, T. 2001. Penanganan Multikolinearitas dengan Analisis Regresi Komponen Utama, *Jurnal Matematika dan Komputer*. **Vol.4 ,No.2 Agustus**

# ANALISIS DISKRIMINAN UNTUK KLASIFIKASI KABUPATEN/KOTA TERTINGGAL DI PROVINSI BENGKULU

Siti Annisa Cholifah<sup>1</sup>, Sigit Nugroho<sup>2</sup>, Pepi Novianti<sup>2</sup>

<sup>1</sup>Mahasiswa Jurusan Matematika, FMIPA UNIB

<sup>2</sup>Dosen Jurusan Matematika, FMIPA UNIB

Jln. WR Supratman Kandang Limun, Bengkulu 38371

**ABSTRACT.** *Purpose of this research is the classification the left regency/city in 2012 and to know the factors influenced the different of left district in Bengkulu Province by using discriminant analysis. Discriminant analysis can separate the regency/city in Bengkulu Province into two group left district and unleft district with calculation the function score of discriminant compare to cut off score. The result of this research is 5 regency are classified as left district, 5 regency are classified as unleft district, and the factors influenced the different of left district is the number of life expectancy and the numbers of midwife, with the percentage of classification of decision is 100%.*

*Key words: Left District, Discriminant Analysis, The Function Score of Discriminant.*

## 1. PENDAHULUAN

Berdasarkan Kementerian Negara Pembangunan Daerah Tertinggal (KNPDT) 2010-2014 yang ditetapkan dengan Perpres No. 5 tahun 2010, pengertian kabupaten/kota tertinggal adalah kabupaten/kota yang masyarakat serta wilayahnya relatif kurang berkembang dibandingkan daerah lain dalam skala nasional. Suatu daerah dikategorikan sebagai daerah tertinggal, disebabkan oleh beberapa faktor, yaitu perekonomian masyarakat, Sumber Daya Manusia, prasarana (infrastruktur), kemampuan keuangan lokal, aksesibilitas, dan karakteristik daerah.

KNPDT 2010-2014 menyatakan bahwa masih tingginya tingkat kesenjangan pembangunan antar wilayah belum optimal, perkembangan kawasan pertumbuhan yang diharapkan menjadi penggerak daerah tertinggal dan kawasan perbatasan. Ini menunjukkan belum adanya keterkaitan dan integrasi ekonomi wilayah dalam sistem pengembangan wilayah.

Provinsi Bengkulu memiliki sepuluh kabupaten/kota dan diantaranya masuk sebagai kabupaten tertinggal. Seperti dijelaskan oleh Staf Khusus Kementerian

Pembangunan Daerah Tertinggal, Jhon Ramadhan, bahwa masih ada enam kabupaten tertinggal yang masuk dalam Kepres tentang Rencana Pembangunan Jangka Menengah Nasional 2010-2014. Meskipun menyatakan ada enam kabupaten tertinggal, namun hanya ada lima kabupaten tertinggal yang disebutkan. Kabupaten tertinggal tersebut adalah Kabupaten Mukomuko, Kabupaten Lebong, Kabupaten Seluma, Kabupaten Kaur dan Kabupaten Kepahiang.

Terdapat berbagai macam metode statistika untuk menganalisis berbagai macam permasalahan, termasuk dalam menentukan status kabupaten/kota tertinggal dan tidak tertinggal. Salah satu metode yang dapat digunakan untuk mengklasifikasikan suatu objek ke dalam kelompok adalah analisis diskriminan. Analisis diskriminan merupakan teknik multivariat yang berkaitan dengan pemisahan objek dalam kelompok yang berbeda dan mengalokasikan objek ke dalam suatu kelompok yang telah ditetapkan sebelumnya.

## **2. TINJAUAN PUSTAKA**

### **2.1 Daerah Tertinggal**

Daerah tertinggal adalah daerah kabupaten/kota yang relatif kurang berkembang dibandingkan daerah lain dalam skala nasional dan berpenduduk relatif tertinggal. Strategi Nasional Pembangunan Daerah Tertinggal menyatakan bahwa ruang lingkup daerah tertinggal adalah wilayah daerah otonom kabupaten/kota. Penetapan status daerah tertinggal mempertimbangkan hal-hal berikut, yaitu: kondisi sosial, ekonomi, budaya, dan wilayah (fungsi inter dan intra spasial baik pada aspek alam, aspek manusia, maupun prasarana pendukung) dari masing-masing daerah. Kondisi tersebut dijadikan dasar dalam menentukan tingkat ketertinggalan suatu daerah dengan menggunakan perhitungan atas kriteria umum, sub kriteria dan indikator yang telah ditentukan bobotnya.

### **2.2 Analisis Diskriminan**

Analisis diskriminan adalah salah satu metode yang dapat digunakan untuk mengetahui variabel mana yang membedakan suatu kelompok dengan kelompok lain dalam suatu populasi. Analisis diskriminan bertujuan mengklasifikasikan suatu objek ke dalam kelompok yang saling lepas (*mutually exclusive/disjoint*) dan menyeluruh (*exhaustive*) berdasarkan sejumlah variabel bebas dan mengelompokkan objek baru ke dalam kelompok-kelompok yang saling bebas tersebut.

Dalam analisis diskriminan ada dua asumsi yang harus dipenuhi (Nugroho, 2008) yaitu:

- a. Memiliki distribusi multivariat normal.
- b. Memiliki kesamaan matriks varians kovarians antar grup.

### 2.3 Model Analisis Diskriminan

Analisis diskriminan termasuk dalam *multivariate dependence method*, dengan model:

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_pX_p$$

$X_p$  = variabel independen, yaitu data berjenis interval atau rasio, seperti usia seseorang, tinggi sebuah pohon, kandungan zat besi dalam tubuh dan sebagainya.

$b_p$  = koefisien fungsi diskriminan yang telah dihasilkan.

$Y$  = variabel dependen, yaitu data kategorikal atau nominal, seperti golongan tidak tertinggal (kode 0), golongan tertinggal (kode 1), dan sebagainya. Jika data kategorikal tersebut hanya terdiri atas dua kode saja, maka model bisa disebut *Two-Group Discriminant Analysis*. Sedangkan jika kode lebih dari dua kategori, disebut dengan *Multiple Discriminant Analysis*.

### 2.4 Analisis Diskriminan Fisher

Prinsip utama dari fungsi diskriminan Fisher adalah pemisahan sebuah populasi. Fungsi diskriminan yang terbentuk dapat digunakan untuk pengelompokkan suatu observasi berdasarkan kelompok-kelompok tertentu. Metode Fisher ini tidak mengasumsikan data harus berdistribusi normal, tetapi dalam perhitungan salah satu syarat yang harus diperhatikan adalah data yang digunakan harus memiliki matriks kovarians yang sama untuk setiap kelompok populasi yang diberikan.

Pada analisis diskriminan dua kelompok, jika variabel bebas itu  $X_1, X_2, \dots, X_p$ , dan kombinasi linier yang dicari adalah  $Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \dots + \beta_pX_p$  maka vektor  $\beta = (\beta_1, \beta_2, \dots, \beta_p)$  perlu dipilih sedemikian hingga hasil bagi  $\frac{\beta' \mu_1 - \beta' \mu_2}{\beta' \Sigma \beta}$  mencapai maksimum dimana  $\beta = c \Sigma^{-1}(\mu_1 - \mu_2)$  untuk sembarang konstan  $c \neq 0$ .

Dimana matriks varians kovarians sampel gabungannya adalah  $S$  dengan rumus:

$$S = \frac{1}{n_1 + n_2 - 2} [(X_1'X_1 + X_2'X_2)]$$

## 2.5 Pengujian Signifikasi

Analisis diskriminan bertujuan juga untuk menghasilkan ukuran tingkat perbedaan antara kelompok-kelompok yang terlibat. Jika fungsi diskriminan untuk dua kelompok  $K_1$  dan  $K_2$  adalah  $Y = b'X$ , maka selisih antara rata-rata skor diskriminan adalah:

$$\begin{aligned}\bar{Y}_1 - \bar{Y}_2 &= b'\bar{X}_1 - b'\bar{X}_2 = b'(\bar{X}_1 - \bar{X}_2) \\ &= (\bar{X}_1 - \bar{X}_2)'S^{-1}(\bar{X}_1 - \bar{X}_2)\end{aligned}$$

Besaran ruas kanan disebut jarak *Mahalanobis* antara  $\bar{X}_1$  dan  $\bar{X}_2$ , dan dinyatakan dengan tanda  $D^2(\bar{X}_1, \bar{X}_2)$  atau  $D^2$ . Ukuran atau statistik  $D^2$  dapat digunakan untuk menyelidiki apakah antara dua vektor rata-rata ada perbedaan yang signifikan.

## 2.6 Klasifikasi

Aturan klasifikasi atau penempatan individu (objek) kedalam kedua kelompok digunakan untuk memprediksi variabel masuk dalam kelompok mana yang kemudian akan didapat matriks klasifikasi. Proses klasifikasi dapat juga digunakan untuk menilai validasi analisis diskriminan.

Aturan klasifikasi menggunakan rumus:

$$Z_{CU} = \frac{n_1 Z_1 + n_2 Z_2}{n_1 + n_2}$$

di mana:

$Z_{CU}$  = angka kritis, yang berfungsi sebagai *cut off score*.

$n_1$  = jumlah sampel di grup 1, yang dalam kasus ini adalah grup tidak tertinggal.

$n_2$  = jumlah sampel di grup 2, yang dalam kasus ini adalah grup tertinggal.

$Z_1$  = angka centroid pada grup 1.

$Z_2$  = angka centroid pada grup 2.

## 3. METODE PENELITIAN

### 3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yaitu data hasil Survei Sosial Ekonomi Nasional (SUSENAS) tahun 2012 yang dilaksanakan oleh Badan Pusat Statistik (BPS). Data yang digunakan berupa data faktor penyebab daerah tertinggal untuk wilayah kabupaten/kota di Provinsi Bengkulu yang terdiri dari 10 kabupaten/kota. Data tersebut disesuaikan dengan ketersediaan data yang ada.

### 3.2 Variabel Penelitian

Variabel-variabel yang digunakan pada penelitian ini yaitu persentase penduduk miskin ( $X_1$ ), angka harapan hidup ( $X_2$ ), jumlah puskesmas ( $X_3$ ), jumlah klinik keluarga berencana ( $X_4$ ), jumlah bidan ( $X_5$ ), angka melek huruf ( $X_6$ ), jumlah sekolah SMP ( $X_7$ ), persentase rumah tangga pengguna listrik ( $X_8$ ).

### 3.3 Analisis Data

Prosedur analisis data adalah sebagai berikut:

1. Tabulasi pada Data View.
2. Uji asumsi dan signifikan.
  - a. *Multivariate normal* dilakukan dengan menguji masing-masing variabel secara individu dengan menggunakan uji *Chi-Square*.
  - b. Kesamaan matriks varian kovarian. Kriteria pengujian untuk kesamaan matriks varian kovarian dengan menggunakan *Box's M*.
  - c. Hasil uji vektor rata-rata dengan menggunakan uji Wilk's Lambda.
3. Pembentukan fungsi diskriminan.
4. Interpretasi hasil dengan cara membandingkan nilai fungsi diskriminan dan  $Z_{CU}$ .
5. Mengukur ketepatan klasifikasi menggunakan rumus *Hit Ratio*.

## 4. HASIL DAN PEMBAHASAN

### 4.1 Uji Asumsi Analisis Diskriminan

#### 1. Uji Normalitas

Uji normalitas pada multivariat bisa dilakukan pada setiap variabel dengan logika bahwa jika secara individual masing-masing variabel memenuhi asumsi normalitas, maka secara bersama-sama (multivariat) variabel-variabel tersebut juga bisa dianggap memenuhi asumsi normalitas.

**Tabel 1 Uji Kenormalan**

Tests of Normality						
	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
X1	.190	10	.200*	.918	10	.344
X2	.194	10	.200*	.947	10	.630
X3	.224	10	.169	.854	10	.065
X4	.188	10	.200*	.890	10	.169
X5	.199	10	.200*	.873	10	.109
X6	.137	10	.200*	.979	10	.957
X7	.230	10	.142	.898	10	.209
X8	.221	10	.182	.955	10	.733

a. Lilliefors Significance Correction  
\*. This is a lower bound of the true significance.

Karena nilai sig. semua variabel pada tabel *kolmogorov-smirnov*  $> 0,05$  maka  $H_0$  diterima. Artinya variabel-variabel independen tersebut berdistribusi normal. Jadi variabel-variabel independen tersebut memenuhi asumsi dan dapat digunakan untuk analisis diskriminan.

## 2. Uji Matriks Varians Kovarians

Pengujian dilakukan untuk menguji varians dari setiap variabel dengan menggunakan uji *Box's M*. Asumsi yang harus terpenuhi adalah *group covariance matrices* setiap variabel relatif sama.

**Tabel 2 Uji Kesamaan Matriks Varians Kovarians**

Test Results		
Box's M		2.030
F	Approx.	.493
	df1	3
	df2	1.152E4
	Sig.	.687

Tests null hypothesis of equal population covariance matrices.

Karena nilai sig. pada tabel *Test Results*  $> 0,05$  maka  $H_0$  diterima. Artinya grup kovarians matrik relatif sama (kedua variabel dependen tertinggal dan tidak tertinggal mempunyai matrik varians-kovarians yang sama pada grup-grup yang ada).

## 3. Uji Vektor Rata-Rata

Untuk melihat kemampuan variabel *independen* dalam membedakan grup secara *multivariate* digunakan uji *Wilks' Lambda*.

**Tabel 3 Uji Vektor Rata-Rata**

Wilks' Lambda				
Test of ...	Wilks' Lambda	Chi-square	df	Sig.
1	.241	9.972	2	.007

Karena nilai sig pada tabel  $0,007 < 0,05$  maka  $H_0$  ditolak. Artinya bahwa ke delapan variabel *independen* mampu membedakan kategori secara signifikan (terdapat perbedaan rata-rata antara dua kategori).

## 4.2 Interpretasi Fungsi Diskriminan

### 1. Pembentukan Fungsi Diskriminan

Dalam membuat fungsi diskriminan, digunakan metode *Stepwise Discriminant Analysis*. Berdasarkan hasil analisis dengan program SPSS diperoleh dua variabel yang membentuk fungsi diskriminan yaitu  $X_2$  dan  $X_5$ , sedangkan variabel  $X_1$ ,  $X_3$ ,  $X_4$ ,  $X_6$ ,  $X_7$ , dan  $X_8$  tidak masuk dalam fungsi diskriminan.

### 2. Fungsi Diskriminan

Menurut Santoso (2002), kegunaan fungsi diskriminan adalah untuk mengetahui *case* (dalam hal ini suatu kabupaten) masuk pada grup yang satu atau pada grup lainnya. Dalam menentukan fungsi diskriminan dapat dilihat pada output SPSS berikut.

**Tabel 4 Koefisien Fungsi Diskriminan**

Canonical Discriminant Function Coefficients	
	Function
	1
x2	.583
x5	.011
(Constant)	-42.503

Unstandardized coefficients

Fungsi diskriminan tersebut dapat ditulis sebagai berikut:

$$Y = -42,503 + 0,583(X_2) + 0,011(X_5) \quad (4.1)$$

Selain fungsi di atas, dengan dipilihnya *Fisher Function Coefficient* pada proses analisis akan terbentuk fungsi diskriminan dari Fisher yang dapat dilihat pada output SPSS.

**Tabel 5 Pengklasifikasian Koefisien Fungsi Diskriminan**

Classification Function Coefficients		
	KATEGORI	
	Tidak Tertinggal	Tertinggal
X2	40.724	38.873
X5	.245	.212
(Constant)	-1455.585	-1320.514

Fisher's linear discriminant functions

Fungsi diskriminan dari Fisher (4.2) hampir sama dengan fungsi diskriminan sebelumnya (4.1), hanya saja pembagiannya berdasarkan kode grup:

- Untuk kabupaten yang termasuk dalam kategori Tidak Tertinggal  
 $Tidak\ Tertinggal = -1455,585 + 40,724(X_2) + 0,245(X_5)$
- Untuk kabupaten yang termasuk dalam kategori Tertinggal  
 $Tertinggal = -1320,514 + 38,873(X_2) + 0,212(X_5)$

Selisih antara grup tidak tertinggal dan tertinggal adalah:

$$Y = -135,075 + 1,851(X_2) + 0,033(X_5) \quad (4.2)$$

Fungsi yang akan digunakan dalam aplikasi pada perhitungan *score* menggunakan fungsi *Unstandardized* (4.1). Sedangkan fungsi Fisher persamaan (4.2) sebenarnya bersifat proporsional dengan fungsi *Unstandardized*, yang untuk kasus ini dengan mengalikan setiap koefisien dari Fisher dengan angka 0,31 maka akan diperoleh fungsi *Unstandardized*.

### 3. Interpretasi Fungsi Diskriminan

**Tabel 6 Function of Group Centroids**

Functions at Group Centroids	
kategori	Function
tidak tertinggal	1.589
tertinggal	-1.589

Unstandardized canonical discriminant functions evaluated at group means

Dari tabel di atas akan dilakukan perhitungan *cut off score* pada persamaan (5) sebagai berikut:

$$Z_{CU} = \frac{n_1 Z_1 + n_2 Z_2}{n_1 + n_2} = \frac{(5 \times 1,589) + (5 \times (-1,589))}{5 + 5} = 0$$

Penggunaan angka  $Z_{CU}$  berdasarkan keputusan:

- Angka skor kasus di atas  $Z_{CU}$ , masuk dalam kategori tidak tertinggal (kode 0).
- Angka skor kasus di bawah  $Z_{CU}$ , masuk dalam kategori tertinggal (kode 1)

Dengan menggunakan fungsi diskriminan *Unstandardized* pada persamaan (4.1) diperoleh hasil sebagai berikut.

Misalkan untuk Kabupaten Bengkulu Selatan, skor fungsi diskriminannya adalah:

$$Y = -42,503 + 0,583(67,77) + 0,011(361) = 0,978$$

Karena *score* diskriminan Kabupaten Bengkulu Selatan adalah  $0,978 > 0$ , maka masuk kedalam kategori tidak tertinggal.

Misalkan untuk Kabupaten Kaur, skor fungsi diskriminannya adalah:

$$Y = -42,503 + 0,583(67,85) + 0,011(92) = -1,934$$

Karena *score* diskriminan Kabupaten Kaur adalah  $-1,934 < 0$ , maka masuk kedalam kategori tertinggal.

#### 4. Ketepatan Hasil Pengklasifikasian

Setelah fungsi diskriminan diperoleh, kemudian melakukan pengujian klasifikasi fungsi diskriminan, selanjutnya akan dilihat seberapa besar ketepatan klasifikasi kategori tersebut. Hal ini dapat dilihat pada output SPSS pada tabel 7 sebagai berikut.

**Tabel 7 Menilai Kelayakan Fungsi Diskriminan**

Classification Results <sup>a,c</sup>					
			Predicted Group Membership		Total
			tidak tertinggal	tertinggal	
Original	Count	tidak tertinggal	5	0	5
		tertinggal	0	5	5
	%	tidak tertinggal	100.0	.0	100.0
		tertinggal	.0	100.0	100.0
Cross-validated <sup>a</sup>	Count	tidak tertinggal	5	0	5
		tertinggal	1	4	5
	%	tidak tertinggal	100.0	.0	100.0
		tertinggal	20.0	80.0	100.0

a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.  
b. 100,0% of original grouped cases correctly classified.  
c. 90,0% of cross-validated grouped cases correctly classified.

Terlihat pada bagian original, kabupaten yang pada data awal berada pada kategori tidak tertinggal dan setelah klasifikasi tetap pada kategori tidak tertinggal sebanyak 5 kabupaten. Sedangkan kabupaten yang pada data awal berada pada kategori tertinggal dan setelah klasifikasi tetap pada kategori tertinggal sebanyak 5 kabupaten.

Hal ini berarti 100% dari 10 kabupaten/kota yang diolah telah dimasukkan pada grup yang sesuai dengan data semula. Jika dilihat dari validasi silang (*cross validated*), pada kode c angka tersebut 90%. Semakin tinggi nilai validasi, termasuk *cross validated groups* tentu semakin bagus, karena semakin tepat fungsi diskriminan membedakan kedua grup.

## 5. KESIMPULAN DAN SARAN

### 1. Kesimpulan

Berdasarkan hasil pengolahan data faktor penyebab daerah tertinggal di Provinsi Bengkulu tahun 2012, dapat diperoleh kesimpulan sebagai berikut:

1. Model fungsi diskriminan yang terbentuk berdasarkan hasil pengolahan data adalah:

$$Y = -42,503 + 0,583(X_2) + 0,011(X_5)$$

2. Faktor yang mempengaruhi perbedaan kabupaten/kota tertinggal di Provinsi Bengkulu yaitu angka harapan hidup dan jumlah bidan.
3. Model diskriminan yang terbentuk ini mempunyai ketepatan pengklasifikasian sebesar 100%. Sehingga model diatas dapat digunakan untuk mengklasifikasikan suatu kabupaten/kota termasuk dalam kategori tertinggal atau tidak tertinggal.

## 2. Saran

Berdasarkan hasil penelitian di atas, saran yang dapat penulis sampaikan untuk pemerintah Provinsi Bengkulu, agar dapat memprioritaskan pembangunan pada daerah yang masih relatif tertinggal dengan memperhatikan faktor sumber daya manusia pada bidang kesehatan yaitu angka harapan hidup dan jumlah bidan. Hal ini dimaksudkan agar masyarakat menjadi sejahtera.

Untuk melakukan analisis diskriminan sebaiknya variabel bebas yang akan dianalisis memiliki pengaruh besar terhadap variabel terikat, sehingga variabel yang layak diuji tidak sedikit. Kemudian, memastikan pengujian asumsi terpenuhi semua agar diperoleh model diskriminan dengan tingkat ketepatan yang lebih tinggi.

## 6. DAFTAR PUSTAKA

Anonim. 2010. *Kementrian Pembangunan Daerah Tertinggal-Penetapan Daerah Tertinggal dalam RPJM 2010-2014*. Jakarta: Kementrian Pembangunan Daerah Tertinggal.

Anonim. 2010. *Peraturan Presiden Republik Indonesia No. 5 Tahun 2010-RPJM 2010-2014*. Jakarta: Kementrian Pembangunan Daerah Tertinggal.

Anonim. 2012. *Enam kabupaten di Bengkulu Masih Tertinggal*. ANTARA Bengkulu: Bengkulu. Publikasi: Kamis, 21 Juni 2012.[Online]. Tersedia dalam <http://www.antarabengkulu.com/berita/4113/enam-kabupaten-di-bengkulu-masih-tertinggal>

Anonim. 2013. *Provinsi Bengkulu dalam Angka 2013*. Bengkulu: Badan Pusat Statistik.

- Johnson, R. A. dan D. W. Wichern. 2002. *Applied Multivariate Statistical Analysis*. Fifth Edition. New York: Prentice-Hall. Inc.
- Maria, T. S. N., D, Dasari., dan N, Herrhyanto. 2012. *Analisis Diskriminan Terhadap Perilaku Mahasiswa dalam Mengkonsumsi Makanan Cepat Saji*. [Online]. Tersedia dalam <http://journal.fpmipa.upi.edu/index.php/eurekamatika/article/view/35> [20 Agustus 2015]
- Nugroho, S. 2008. *Statistika Multivariat Terapan*. Unib Press. Bengkulu.
- Santoso, S. 2002. *Buku Latihan SPSS Statistik Multivariat*. Jakarta: PT Elex Media Komtindo.
- Sulistiyorini, I. 2013. *Analisis Faktor-Faktor yang Mempengaruhi Kepuasan Pelanggan Internet Speedy Reguler Menggunakan Analisis Diskriminan di Semarang*. [Skripsi]. Jurusan Matematika. Fakultas Matematika dan Ilmu Pengetahuan Alam. Universitas Negeri Semarang.
- Suryanto. 1988. *Metode Statistika Multivariat*. Jakarta: Departemen Pendidikan dan Kebudayaan Direktorat Jendral Pendidikan Tinggi, Proyek Pengembangan Lembaga Tinggi Tenaga Kependidikan.

# ANALISIS INDEKS HARGA SAHAM GABUNGAN DENGAN PENDEKATAN VECTOR ERROR CORRECTION MODEL (VECM) (Studi Kasus: Data IHSG Periode Juli 2005 s/d Maret 2016)

Wisnu Setia Nugroho<sup>1)</sup>, Sigit Nugroho<sup>2)</sup>, Jose Rizal<sup>2)</sup>

<sup>1)</sup> Alumni Jurusan Matematika FMIPA Universitas Bengkulu

<sup>2)</sup> Staf Pengajar Jurusan Matematika FMIPA Universitas Bengkulu

---

## ABSTRAK

Penelitian ini bertujuan untuk menetapkan pemodelan Indeks Harga Saham Gabungan (IHSG) dengan *Vector Error Correction Model* (VECM) berdasarkan periode Juli 2005 sampai dengan Maret 2016 dan meramalkan IHSG untuk 3 bulan ke depan. Dalam pemodelan digunakan variabel yang mempengaruhi IHSG yaitu (i) *BI Rate*, (ii) Indeks Saham Global (S&P 500), (iii) Inflasi, (iv) Jumlah Uang Beredar (M1), dan (v) Kurs Rupiah. Penelitian diawali dengan melihat kestasioneran data dengan menggunakan uji *Phillip-Perron* (PP), bila data belum stationer aka dilakukan tahap *differencing* untuk menstationerkan data. Kemudian dilakukan penentuan *lag* optimal dengan melihat nilai *Aikake Information Criteria* (AIC) dan *Schwarz Information Criteria* (SIC), selanjutnya dilakukan uji kontegrasi untuk mengetahui apakah terdapat hubungan jangka panjang antar variabel terhadap IHSG sebagai syarat untuk melanjutkan ke pemodelan VECM. Kemudian dilakukan pemodelan IHSG dengan pendekatan VECM. Untuk melihat perilaku dinamis dari model VECM melalui *Impulse Respose Function* (IRF) dan *Variance Decomposition* (VD). Hasil Peramalan menggunakan model VECM untuk 3 periode ke depan menghasilkan nilai ramalan pada bulan April 2016 sebesar 4965.095, bulan Mei 2016 sebesar 4997.200, dan bulan Juni 2016 sebesar 5029.304.

**Kata Kunci :** IHSG, VECM, IRF, VD, PP

---

## I. PENDAHULUAN

Pasar modal memiliki peran penting bagi perekonomian suatu negara karena pasar modal menjalankan dua fungsi, yaitu pertama sebagai sarana bagi pendanaan usaha atau sebagai sarana bagi perusahaan untuk mendapatkan dana dari masyarakat pemodal (investor). Dana yang diperoleh dari pasar modal dapat digunakan untuk pengembangan usaha, ekspansi, penambahan modal kerja dan lain-lain, kedua pasar modal menjadi sarana bagi masyarakat untuk berinvestasi pada instrument keuangan seperti saham, obligasi, reksa dana, dan lain-lain (Anonim, 2010).

Saham (*stock*) merupakan salah satu instrumen pasar keuangan yang paling populer. Saham dapat didefinisikan sebagai tanda persertaan modal seseorang atau pihak (badan usaha) dalam suatu perusahaan atau perseroan terbatas. Jika modal tersebut disertakan dalam saham maka pihak badan usaha memiliki klaim atas pendapatan perusahaan, klaim atas aset perusahaan, dan berhak hadir dalam Rapat Umum Pemegang Saham (RUPS). Indeks merupakan salah satu pedoman bagi investor untuk melakukan investasi di pasar modal, khususnya saham. Indeks harga saham adalah indikator atau cerminan pergerakan harga saham (Anonim, 2010).

Indeks Harga Saham Gabungan (IHSG) merupakan indeks gabungan dari seluruh jenis saham Indonesia yang tercatat di bursa efek. IHSG berubah setiap hari karena perubahan harga pasar yang terjadi setiap hari dan adanya saham tambahan.

Banyak teori dan penelitian terdahulu yang mengungkapkan bahwa pergerakan Indeks Harga Saham Gabungan dipengaruhi oleh beberapa faktor.

Menurut Purba (2014), perkembangan pasar modal pada IHSG di Indonesia dipengaruhi oleh berbagai faktor, baik faktor internal dan eksternal. Tingkat pertumbuhan ekonomi yang berlangsung cukup tinggi, tidaklah secara otomatis mengakibatkan membaiknya situasi pasar modal. Pergerakan IHSG yang dipengaruhi oleh berbagai faktor sulit diprediksi dan membutuhkan analisis ekonometrika serta nalar yang tepat untuk menentukan keputusan investasi dalam perusahaan.

Ada banyak metode yang digunakan dalam analisis ekonometrika. Analisis tersebut bertujuan untuk memodelkan hubungan antara variabel-variabel. Variabel tersebut dapat berupa variabel deret waktu. Beberapa pendekatan yang dapat digunakan dalam analisis deret waktu dengan beberapa variabel (multivariat) yaitu, analisis *vector autoregressive* (VAR) dan *vector error correction model* (VECM).

Analisis *Vector Autoregressive* (VAR) merupakan analisis peramalan deret waktu multivariat yang paling sering digunakan untuk memprediksi data stationer. Menurut Sim (1980) VAR biasanya digunakan untuk menganalisis hubungan antara variabel-variabel deret waktu, menganalisis hubungan sistem variabel-variabel deret waktu dan menganalisis dampak dinamis dari faktor gangguan yang terdapat dalam sistem variabel

tersebut. Ketika data tidak stationer dan terdapat hubungan kointegrasi maka VAR terestriksi yang digunakan untuk menganalisis data membuat model yang mempunyai hubungan jangka panjang dan pendek terhadap variabel yang mempunyai masalah non stasioner disebut juga VECM (Sari, 2015).

VECM merupakan bentuk multivariat dari *Error Correction Model* (ECM). ECM adalah teknik untuk mengoreksi ketidakseimbangan jangka pendek menuju keseimbangan jangka panjang, serta dapat menjelaskan hubungan antara peubah terikat dengan peubah bebas pada waktu sekarang dan lampau. Pada penelitian ini difokuskan dengan beberapa variabel (multivariat) deret waktu sehingga penggunaan ECM tidak sesuai karena bentuknya berupa univariat.

Analisis VECM merupakan analisis yang dikembangkan oleh Engle dan Granger (1987) untuk melakukan rekonsiliasi perilaku ekonomi jangka pendek dengan variabel ekonomi jangka panjang (Gujarati, 2004). Konsep penting dalam VECM adalah keseimbangan jangka panjang dari data deret waktu yang sering disebut kointegrasi. Kointegrasi merupakan kombinasi linier dari variabel yang tidak stasioner dan terintegrasi pada ordo yang sama. Kointegrasi bertujuan untuk mengetahui hubungan keseimbangan dalam jangka panjang antara peubah-peubah yang diamati (Enders, 2004). Analisis VECM juga disebut sebagai model VAR restriksi dengan data tidak stationer namun terdapat hubungan kointegrasi.

IHSG yang merupakan salah satu pedoman bagi investor untuk melakukan investasi di pasar modal menjadi barometer ekonomi suatu negara. Apakah kondisi perekonomian dalam negara sedang dalam keadaan baik ataupun buruk sekaligus sebagai penunjang ekonomi negara. Beberapa penelitian terdahulu mengenai IHSG dapat diketahui bahwa faktor eksternal dan internal dapat mempengaruhi IHSG. Heriyanto dan Chen (2014) dalam penelitiannya data IHSG serta faktor yang mempengaruhi yaitu kurs rupiah, jumlah uang beredar, indeks harga konsumen dan indeks S&P 500 mengatakan bahwa terdapat adanya hubungan keseimbangan jangka panjang antar variabel. Yanuar (2013) juga mengatakan bahwa pergerakan dan perubahan pada faktor internal dan eksternal berdampak positif maupun negatif terhadap IHSG. Purba (2014) juga memperoleh hasil bahwa IHSG dipengaruhi oleh variabel indeks harga pedagang besar dan jumlah uang beredar. Berdasarkan beberapa penelitian tersebut memperlihatkan bahwa data yang digunakan merupakan data yang tidak stasioner.

Berdasarkan uraian diatas, maka Peneliti tertarik membahas analisis IHSG dengan pendekatan VECM.

## II. TINJAUAN PUSTAKA

### A. Indeks Harga Saham Gabungan (IHSG)

#### 1. Pengertian IHSG

Indeks adalah ukuran statistik yang digunakan untuk menyatakan perubahan-perubahan perbandingan nilai suatu variabel tunggal atau nilai sekelompok variabel yang berhubungan satu sama lain. Indeks harga adalah suatu angka yang digunakan untuk melihat perubahan mengenai harga dalam waktu dan tempat yang sama ataupun berlainan (Amrullah, 2013). Indeks harga dapat dipakai untuk harga satu saham, sekumpulan saham dalam suatu industri tertentu atau semua saham gabungan (IHSG). Indeks harga saham merupakan angka yang tersusun dengan hitungan tertentu sehingga menghasilkan trend. Indeks Harga Saham Gabungan (IHSG) di mana semua saham tercatat sebagai komponen penghitungan indeks.

#### 2. Metode Perhitungan IHSG

IHSG pada BEI (Bursa efek indonesia) dihitung menggunakan metodologi rata-rata tertimbang berdasarkan jumlah saham tercatat (nilai pasar) atau *Market Value Weighted Average Index*. Menurut Halim (2005) dalam Wijaya (2013), Indeks Harga Saham Gabungan (IHSG) dapat dihitung dengan menggunakan rumus berikut:

$$IHSG_t = \frac{NP_t}{ND} \times 100 \quad (2.1.1)$$

Keterangan:

$IHSG_t$  : Indeks Harga Saham Gabungan pada hari ke- $t$

$NP_t$  : Nilai pasar pada hari ke- $t$ , dari jumlah saham yang tercatat di bursa dikalikan dengan harga pasar per lembar

$ND$  : Nilai dasar

Untuk menghitung rata-rata IHSG adalah dengan membagi Jumlah IHSG periode selama 1 bulan dengan jumlah periode waktu selama 1 bulan.

$$\text{Rata - rata IHSG} = \frac{\text{Jumlah IHSG periode selama 1 bulan}}{\text{Jumlah periode waktu selama 1 bulan}} \quad (2.1.2)$$

#### 3. Faktor Internal dan Eksternal Pengaruh IHSG

Pergerakan Indeks Harga Saham Gabungan dipengaruhi oleh beberapa faktor, seperti faktor eksternal dan faktor internal. Faktor eksternal tersebut bisa datang dari Indeks bursa asing negara lain (*Dow Jones, NIKKEI, NYSE, S&P 500*), pergerakan harga minyak mentah, harga emas dunia dan adanya peristiwa yang mempengaruhi keamanan dan stabilitas perekonomian dunia (seperti kerusakan maupun serangan teroris). Sedangkan faktor internal bisa datang dari kurs mata uang, tingkat pendapatan nasional, jumlah uang beredar, tingkat inflasi, dan suku bunga yang terjadi di negara tersebut (Divianto, 2013).

Berdasarkan pemaparan beberapa faktor yang mempengaruhi IHSG di atas maka penelitian ini mengambil beberapa faktor eksternal dan faktor internal sebagai faktor yang mempengaruhi pergerakan IHSG. Beberapa faktor tersebut yaitu, “Data Tingkat Inflasi, Jumlah Uang Beredar, BI Rate, Indeks Bursa Saham Global (Indeks S&P 500), Serta Nilai Tukar Rupiah”.

## B. Analisis Deret Waktu

Deret waktu adalah suatu himpunan pengamatan yang diperoleh pada titik waktu yang berbeda dengan selang waktu yang sama dan barisan data diasumsikan saling berhubungan satu sama lain (Box and Jenkins, 1994).

### 1. Stationeritas

Data deret waktu merupakan sekumpulan nilai suatu variabel yang diambil pada waktu yang berbeda. Setiap data dikumpulkan secara berkala pada interval waktu tertentu. Juanda dan Junaidi (2012) menyatakan data deret waktu dikatakan stasioner jika memenuhi tiga kriteria, yaitu nilai tengah (rata-rata) dan ragamnya konstan dari waktu ke waktu, serta kovarian antara dua data runtun waktu hanya tergantung dari lag antara dua periode waktu tersebut. Secara statistik dinyatakan sebagai berikut:

$$1. E(Y_t) = \mu \text{ rata-rata } Y \text{ konstan} \quad (2.2.1)$$

$$2. Var(Y_t) = E(Y_t - \mu)^2 \text{ ragam } Y \text{ konstan} \quad (2.2.2)$$

$$3. \gamma_k = E[(Y_t - \mu)(Y_{t+k} - \mu)] \text{ kovarian} \quad (2.2.3)$$

Berdasarkan nilai tengah dan ragamnya, terdapat dua jenis kestasioneran data yaitu stasioner dalam rata-rata dan stasioner dalam ragam. Data dikatakan stasioner dalam rata-rata jika rata-rata cenderung konstan dari waktu ke waktu atau data bersifat stabil. Dan data deret waktu dikatakan stasioner dalam ragam jika fluktuasi datanya tetap atau konstan. Stasioneritas berarti bahwa tidak terjadinya pertumbuhan dan penurunan data. Data deret waktu sangat banyak digunakan, namun data deret waktu menyimpan berbagai permasalahan yaitu salah satunya adalah autokorelasi yang mengakibatkan data menjadi tidak stasioner. Ada beberapa cara untuk melihat kestasioneran data, yaitu dengan melihat koefisien autokorelasi dan korelogramnya (*correlogram*) serta dengan uji akar unit.

Uji stasioneritas dengan Uji Akar Unit Phillips-Perron (PP), dilakukan dengan hipotesis :

$$H_0: \delta = 0 \text{ (d}Y_t \text{ stasioner)}$$

$$H_1: \delta \neq 0 \text{ (}Y_t \text{ tidak stasioner)}$$

Statistik Uji untuk Uji PP adalah:

$$t_{pp} = \frac{(\delta - 1)}{SE(\delta)} \quad (2.2.8)$$

$$p_{value} = Z(t_{pp}) \quad (2.2.9)$$

$H_0$  ditolak jika  $|t| < |t_{McKinnon}|$  dan disimpulkan bahwa  $Y_t$  mengandung akar unit atau  $Y_t$  tidak stasioner. Sedangkan  $H_0$  diterima jika  $|t| >$

$|t_{McKinnon}|$  dan disimpulkan bahwa  $Y_t$  tidak mengandung akar unit atau  $Y_t$  stasioner.

Apabila data tidak stasioner pada rata-rata maka perlu dilakukan *differencing* (pembedaan). *Differencing* merupakan pengurangan data tertentu dengan data sebelumnya.

### C. Vector Error Correction Model (VECM)

VECM dipopulerkan pertama kali oleh Engle dan Granger dalam mengoreksi *disequilibrium* jangka pendek terhadap jangka panjangnya. VECM digunakan untuk mengestimasi data yang tidak stasioner pada tingkat level, namun memiliki hubungan kointegrasi. Model ini pada dasarnya menggunakan bentuk VAR yang terestriksi. Restriksi tambahan ini harus diberikan karena keberadaan bentuk data yang tidak stasioner namun terkointegrasi. VECM kemudian memanfaatkan informasi restriksi kointegrasi tersebut ke dalam spesifikasi model. Spesifikasi ini meretriksi hubungan jangka panjang variabel-variabel endogen agar konvergen ke dalam hubungan kointegrasinya, namun tetap membiarkan keadaan dinamis jangka pendek. Sehingga VECM juga sering disebut sebagai model VAR bagi data deret waktu yang bersifat non stasioner dan memiliki hubungan kointegrasi sehingga disebut sebagai VAR yang terestriksi (Prakoso, 2009).

Jika suatu data deret waktu model VAR terbukti terdapat hubungan kointegrasi, maka VECM dapat digunakan untuk mengetahui tingkah laku jangka pendek dari suatu variabel terhadap nilai jangka panjangnya. VECM adalah model untuk menganalisis data *multivariate time series* yang tidak stasioner. Secara umum model VAR yang tidak terestriksi dan memiliki sampai  $p$ -lags adalah sebagai berikut:

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + \varepsilon_t \quad (2.3.1)$$

keterangan,

$Y_t$  : sebuah vektor dengan  $k$  variabel

$A$  : parameter matriks

$\varepsilon_t$  : vektor error

Karena adanya hubungan kointegrasi secara linear maka Persamaan (2.3.1) model VAR akan berubah menjadi model VECM dengan menggunakan  $y_{t-1}$  (*first difference*), yaitu:

$$\Delta y_t = \Pi y_{t-1} + \sum_{i=1}^k \Gamma_i \Delta y_{t-i} + \varepsilon_t \quad (2.3.2)$$

dengan

$$\Pi = -(1_k - A_1 - \dots - A_p), \text{ dan } \Gamma_i = (A_{i+1} + \dots + A_p), i = 1, \dots, p - 1$$

Keterangan,

$\Gamma_j$  : koefisien matriks ( $p \times p$ ),  $j = 1, \dots, k$

$\Pi$  : matriks ( $p \times r$ );  $0 < r < p$  dan  $r$  merupakan jumlah kombinasi linier elemen  $y_t$  yang hanya dipengaruhi oleh *shock transistor*.

$\mu$  : vektor *error correction*

$t$  : jumlah observasi

#### D. Penentuan Panjang Lag

Secara umum terdapat beberapa parameter yang dapat digunakan untuk menentukan panjang lag yang optimal, antara lain AIC (*Akaike Information Criterion*) dan SIC (*Schwarz Information Criterion*). Penentuan panjang lag yang optimal didapat dari persamaan VAR dengan nilai AIC, atau SIC yang terkecil.

Sebagaimana dinyatakan Enders (2004) perhitungan dari AIC dan SIC adalah sebagai berikut:

$$AIC(k) = T \ln \left( \frac{SSR(k)}{T} \right) + 2n \quad (2.4.1)$$

$$SIC(k) = T \ln \left( \frac{SSR(k)}{T} \right) + n \ln(T) \quad (2.4.2)$$

Keterangan:

$k$  : panjang lag

$SSR$  : jumlah kuadrat residual

$n$  : jumlah parameter yang diestimasi

$T$  : jumlah observasi yang digunakan

Nilai  $k$  untuk panjang lag ditentukan terlebih dahulu dari persamaan VAR yang stabil hingga didapat lag maksimum yang dihasilkan oleh sistem VAR sebagai nilai  $k$  yang digunakan.

#### E. Uji Kointegrasi Johansen

Prosedur Johansen memiliki peluang untuk menguji bentuk vektor kointegrasi yang terbatas. Untuk menguji batasan pada vektor kointegrasi, Johansen menentukan dua matriks  $\alpha$  dan  $\beta$ , keduanya dengan dimensi ( $p \times r$ ), dimana  $r$  merupakan peringkat dari  $n$ , sehingga:

$$\Pi = \alpha\beta' \quad (2.5.1)$$

dengan,

$\alpha$  : bobot matriks dari setiap vektor kointegrasi dengan ukuran  $n \times r$ .

$\beta$  : matriks parameter kointegrasi berukuran  $r \times n$

Pendefinisian suatu vektor pada pendekatan Johansen diawali dengan vektor dari  $n$  potensial variabel endogen  $Y_t$ .  $Y_t$  diasumsikan sebagai suatu sistem VAR yang tidak terestriksi dan memiliki lag sampai  $k$ -lags:

$$Y_t = A_1 Y_{t-1} + \dots + A_k Y_{t-k} + \varepsilon_t \quad (2.5.2)$$

keterangan,

$Y_t$  : sebuah vektor dengan  $k$  variabel non stationer  $I(1)$

$A$  : parameter matriks

$\varepsilon_t$  : vektor error

Persamaan (2.5.2) dapat diformulasikan kembali ke dalam bentuk *Vector Error Correction Model* (VECM) dengan menggunakan  $Y_{t-1}$  (*first difference*) sebagai berikut:

$$\Delta Y_t = \Gamma_1 \Delta Y_{t-1} + \dots + \Gamma_{k-1} \Delta Y_{t-k+1} + \Pi Y_{t-k} + \varepsilon_t \quad (2.5.3)$$

dengan

$$\Pi = \sum_{i=1}^p A_i - P, \text{ dan } \Gamma_i = -\sum_{j=i+1}^p A_j, i = 1, \dots, p-1$$

(Anonim, 2013).

Untuk pengujian hipotesis dapat digunakan Uji *trace statistic*

$$LR_{tr}(r|k) = -T \sum_{i=r+1}^k \log(1 - \lambda_i)$$

Uji *maximum eigenvalue*

$$LR_{max}(r|r+1) = -T \log(1 - \lambda_i) \\ = LR_{tr}(r|k) - LR_{tr}(r+1|k)$$

untuk  $r = 0, 1, \dots, k-1$ .

Tolak  $H_0$  jika jika statistik uji trace dan nilai Eigen maksimum  $>$  nilai kritis pada saat  $\alpha$ , atau  $p$  value  $<$  nilai signifikansi  $\alpha$ , Pada tingkat signifikansi  $(1 - \alpha)100\%$ .

#### F. Uji Causality Granger

Uji *Granger Causality* digunakan untuk melihat pengaruh masing-masing variabel terhadap variabel lainnya satu per satu. Berdasarkan pada hipotesis kausalitas Granger (Subagyo, 2013), uji kausalitas dilakukan untuk mengetahui apakah suatu variabel endogen dapat diperlakukan sebagai variabel eksogen. Hal ini bermula dari ketidaktahuan keterpengaruhannya antar variabel. Jika ada dua variabel  $y$  dan  $z$ , maka apakah  $y$  menyebabkan  $z$  atau  $z$  menyebabkan  $y$  atau berlaku keduanya atau tidak ada hubungan keduanya. Variabel  $y$  menyebabkan variabel  $z$  artinya berapa banyak nilai  $z$  pada periode sekarang dapat dijelaskan oleh nilai  $z$  pada periode sebelumnya dan nilai  $y$  pada periode sebelumnya.

#### G. IRF dan VD

Perilaku dinamis dari model VECM dapat dilihat melalui respon dari setiap variabel terhadap *shock* dari variabel tersebut maupun terhadap variabel endogen lainnya. Terdapat dua cara untuk melihat karakteristik dinamis model VECM, yaitu melalui *Impulse Response Function* dan *Variance Decomposition*. IRF berfungsi untuk mengukur besaran (perubahan dalam persen), orientasi (meningkat atau menurun), dan panjang (seberapa lama *shock* mempengaruhi variabel-variabel terikat) dari suatu respon dan mengevaluasi kecepatan dari mekanisme transmisi dalam beroperasi. *Variance Decomposition* berguna untuk mengukur perkiraan varians error suatu variabel yaitu seberapa besar kemampuan suatu variabel dalam memberikan penjelasan pada variabel lainnya atau pada variabel itu sendiri.

#### H. Peramalan

Seperti analisis peramalan pada umumnya, untuk menentukan keakuratan hasil ramalan dari sebuah model dapat menggunakan *Mean Absolute Error* (MAE) yaitu rata-rata nilai *Absolute Error* dari kesalahan meramal tidak dihiraukan tanda

positif ataupun negatifnya dengan formula sebagai berikut:

$$MAE = \frac{\sum |X_t - \hat{X}_t|}{n} \quad (2.9.1)$$

Keterangan :

$X_t$  = data sebenarnya

$\hat{X}_t$  = data ramalan dihitung dari model yang digunakan waktu t

n = banyaknya data hasil ramalan

### III. METODE PENELITIAN

#### A. Jenis Penelitian

Penelitian ini merupakan penelitian statistik terapan, yaitu suatu penelitian yang dilakukan dengan pengaplikasian teori ke dalam bidang tertentu.

#### B. Sumber Data Penelitian

Data-data diperoleh dari beberapa sumber yaitu, data tingkat Inflasi, Jumlah Uang Beredar, BI Rate, serta nilai tukar rupiah bersumber dari (<http://www.bi.go.id/web/id>), data Indeks Bursa Saham Global (Indeks S&P 500) dan Indeks Harga Saham Gabungan bersumber dari (<http://www.finace.yahoo.com>).

#### C. Variabel Penelitian

Variabel yang digunakan pada penelitian ini adalah deret waktu (*time series*) yaitu data tingkat Inflasi, Jumlah Uang Beredar, BIRate, Indeks Bursa Saham Global, nilai tukar rupiah dan Indeks Harga Saham Gabungan (IHSG). Data yang digunakan merupakan data bulanan dengan periode Juli 2005 s/d Maret 2016 yang berjumlah 129 observasi.

#### D. Analisis Data

Proses analisis data dilakukan dengan menggunakan bantuan *software Eviews* guna memudahkan Peneliti dalam analisis. Adapun tahapan analisis yang dilakukan adalah sebagai berikut:

1. Plot data untuk melihat kestasioneran data dengan kolelogram pada ACF dan PACF.
2. Pengujian kestasioneran data. Pada pengujian stationeritas data digunakan pengujian *Philip Perron*. Data pada metode VECM berupa data tidak stationer pada tingkat level setelah dilihat dari plot data sehingga dilakukan *differencing* tahap pertama untuk menstationerkan data hingga data stationer pada tingkat yang sama.
3. Jika data sudah stationer pada tingkat yang sama dilanjutkan dengan penentuan panjang *lag* optimum dengan menggunakan AIC dan SIC.
4. Kemudian dilakukan pengujian kointegrasi dengan menggunakan metode *Johansen Cointegration Test* untuk dapat dilanjutkan penaksiran untuk model VECM.
5. Penaksiran parameter model VECM.

6. Uji diagnosis Model VECM menggunakan Uji Portmanteau.
7. Analisis IRF dan FEDV
8. Pengujian *Granger Causality* untuk melihat hubungan antar variabel.
9. Peramalan dengan model VECM.

## IV. PEMBAHASAN

### A. Deskripsi Data

Perhitungan statistik deskriptif dilakukan dengan bantuan *software Eviews 9* yang menghasilkan output pada Tabel 1. sebagai berikut:

**Tabel 1.** Tabel Statistik Data Penelitian

	IHSG	INFLASI	BI_RATE	KURS	S_P	M1
Mean	3244.054	0.071	0.077	10235.26	1443.419	623822.6
Median	3501.300	0.063	0.075	9485.00	1377.940	571352.0
Maximum	5518.675	0.183	0.127	14657.00	2107.390	1064664.
Minimum	1050.090	0.024	0.057	8508.00	735.090	266870.0
Std. Dev.	1354.650	0.035	0.018	1558.29	346.673	241535.7
Skewness	-0.097	1.468	1.329	1.104	0.435	0.262
Kurtosis	1.626	4.707	4.250	2.964	2.355	1.770
Jarque-Bera	10.341	62.059	46.397	26.230	6.305	9.606
Probability	0.005681	0.000000	0.000000	0.000002	0.042	0.008
Observations	129	129	129	129	129	129

Berdasarkan Tabel 1. diperoleh: (1) nilai mean, median, max, min dan Std. Dev. yang dapat dilihat pada tabel; (2) data tidak menyebar Normal, hal ini dapat dilihat berdasarkan nilai probability dari statistik uji Jarque Berra sebesar 0,00 yang lebih kecil jika dibandingkan dengan taraf nyata pengujian sebesar 5% atau 0.05; (3) tidak terdapat data hilang, hal ini terlihat dari besar observasi yang sama untuk setiap variabel.

### B. Pengujian Stationeritas

Langkah awal dalam prosedur ini adalah melakukan uji akar unit terhadap ketiga data tersebut. Berdasarkan pengolahan data, diperoleh hasil seperti yang diperlihatkan pada Tabel 2.

**Tabel 2.** Uji Akar Unit

Data	Nilai Kritis ( $\alpha$ )	Level		First Difference	
		PP Stat	p value	PP Stat	p value
IHSG		-1.0955	0.7163	-9.8096	0,0000
	5%	-2.8841		-2.8842	
S&P		-0.3215	0.9174	-10.1526	0,0000
	5%	-2.8841		-2.8842	
JUB		1.2564	0.9984	-16.0269	0,0000
	5%	-2.8841		-2.8842	
KURS		-0.6015	0.8652	-10.6201	0,0000
	5%	-2.8841		-2.8842	
INFLASI		-2.3623	0.1545	-9.0858	0,0000
	5%	-2.8841		-2.8842	
BI_RATE		-1.5962	0.4816	-3.9019	0.0027
	5%	-2.8841		-2.8842	

Berdasarkan Tabel 2. menunjukkan bahwa data pada tingkat level semuanya tidak stationer dilihat dari nilai p-value > nilai  $\alpha$  pada taraf nyata 1%,5% maupun 10%. Sehingga selanjutnya dilakukan *differencing* agar data menjadi stationer. Setelah itu kembali dilakukan pengujian akar unit *Phillips perron*, pada tabel tersebut diperoleh nilai statistik

nilai p-value < nilai  $\alpha$  pada taraf nyata 1%,5% maupun 10%. Hal ini memberikan cukup bukti bahwa data hasil *differencing* 1 tidak lagi mengandung akar unit atau data telah stasioner sehingga dapat dilanjutkan pada proses identifikasi model.

### C. Pengujian Panjang Lag

Berdasarkan melihat nilai AIC (*Akaike Information Criterion*) dan SIC (*Schwarz Information Criterion*) untuk menentukan panjang lag optimalnya, terlihat bahwa dengan kriteria AIC lag optimal yang disarankan adalah 11 karena berdasarkan pada nilai AIC terkecil, sedangkan kriteria SIC menyarankan lag optimal adalah lag 1 berdasarkan pada nilai SIC terkecil. Untuk melihat model terbaik adalah dengan melihat nilai *Adj.R-squared* pada sistem VAR, dimana perbandingan tersebut pada nilai lag 11 *Adj.R-squared* = 0.984089 dan lag 1 *Adj.R-squared* = 0.984511. Diperoleh nilai lag optimal adalah lag 1 dimana menjadi lag optimal pada VAR, untuk lag optimal adalah 1 dimana VECM (1).

### D. Penentuan Kointegrasi

**Tabel 3.** Uji Kointegrasi Johansen

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.302207	130.8455	117.7082	0.0057
At most 1	0.178897	85.14662	88.80380	0.0892
At most 2	0.163651	60.11410	63.87610	0.0995
At most 3	0.139464	37.41801	42.91525	0.1591
At most 4	0.118325	18.34269	25.87211	0.3214
At most 5	0.018328	2.349286	12.51798	0.9433

Trace test indicates 1 cointegrating eqn(s) at the 0.05 level

\* denotes rejection of the hypothesis at the 0.05 level

\*\*MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	0.302207	45.69884	44.49720	0.0368
At most 1	0.178897	25.03253	38.33101	0.6694
At most 2	0.163651	22.69609	32.11832	0.4405
At most 3	0.139464	19.07532	25.82321	0.3002
At most 4	0.118325	15.99341	19.38704	0.1455
At most 5	0.018328	2.349286	12.51798	0.9433

Max-eigenvalue test indicates 1 cointegrating eqn(s) at the 0.05 level

\* denotes rejection of the hypothesis at the 0.05 level

\*\*MacKinnon-Haug-Michelis (1999) p-values

Hasil uji kointegrasi untuk lag 1 Berdasarkan Tabel 3. di atas dapat dilihat bahwa nilai *trace statistic* dan *maximum eigenvalue* pada  $r = 1$  lebih besar dari *critical value* dengan taraf signifikansi

5%. Hal ini berarti hipotesis nol yang menyatakan bahwa tidak ada kointegrasi ditolak dan hipotesis alternatif yang menyatakan bahwa ada kointegrasi tidak dapat ditolak. Sehingga dapat dilihat bahwa di antara keenam variabel dalam penelitian ini, terdapat minimal satu kointegrasi pada tingkat signifikansi 5%. Dengan demikian, dari hasil uji kointegrasi mengindikasikan bahwa di antara memiliki hubungan stabilitas/ keseimbangan dan kesamaan pergerakan dalam jangka panjang. Dengan kalimat lain, dalam setiap periode jangka pendek, seluruh variabel cenderung saling menyesuaikan, untuk mencapai ekuilibrium jangka panjangnya. Bentuk vektor residual dari persamaan kointegrasi adalah :

$$\begin{aligned} IHSg_t = & 0.016 * (IHSg(-1) - 18041.633 \\ & * INFLASI(-1) + 21491.578 \\ & * BI\_RATE(-1) + 0.373 \\ & * KURS(-1) - 1.554 * S\_P(-1) \\ & + 0.002 * M1(-1) - 50.076 \\ & * @TREND(05M07) - 3176.334 ) \end{aligned}$$

Berdasarkan hasil yang diperoleh maka dapat dilakukan estimasi bentuk persamaan VECM.

### E. Estimasi Model VECM

Setelah didapati hubungan kointegrasi di antara keenam variabel penelitian, maka tahap selanjutnya adalah membentuk model VECM. Menurut Enders, jika terdapat hubungan kointegrasi di antara variabel penelitian, maka estimasi dilakukan dengan VECM, sedangkan jika tidak ada kointegrasi di antara kelima variabel di atas maka estimasi dilakukan dengan VAR-d. Prosedur penting dalam mengestimasi persamaan VECM adalah pemilihan lag optimum. Prosedur pemilihan lag optimum dalam VECM dapat menggunakan kriteria informasi, yaitu AIC dan SIC. Pada pengujian panjang lag diperoleh bahwa lag optimum yang diperoleh adalah lag 1 maka bentuk persamaan VECM yang diestimasi adalah VECM(1).

Dengan demikian, bentuk persamaaan VECM(1) untuk variabel dependen IHSg adalah sebagai berikut:

$$\begin{aligned} D(IHSg) = & 0.016 * (IHSg(-1) - 18041.633 \\ & * INFLASI(-1) + 21491.578 \\ & * BI\_RATE(-1) + 0.373 \\ & * KURS(-1) - 1.554 * S\_P(-1) \\ & + 0.002 * M1(-1) - 50.076 \\ & * @TREND(05M07) - 3176.334 ) \\ & + 0.058 * D(IHSg(-1)) \\ & - 1253.786 * D(INFLASI(-1)) \\ & - 6295.125 * D(BI\_RATE(-1)) \\ & + 0.029 * D(KURS(-1)) + 0.285 \\ & * D(S\_P(-1)) - 0.0009 \\ & * D(M1(-1)) + 30.409 \end{aligned}$$

## F. Uji Diagnosis

Setelah diperoleh model kemudian dilakukan pemeriksaan model yang dilakukan adalah uji asumsi residual dari model tersebut, yaitu uji *portmanteau*.

**Tabel 4.6.1.** Uji Diagnosis *Portmanteau*

Lags	Q-Stat	Prob.	Adj Q-Stat	Prob.	df
1	7.851377	NA*	7.913690	NA*	NA*
2	57.91511	0.7779	58.77844	0.7528	67
3	78.26817	0.9669	79.62391	0.9576	103
4	111.7636	0.9567	114.2087	0.9388	139
5	148.6401	0.9266	152.5964	0.8882	175
6	188.1026	0.8697	194.0158	0.7931	211
7	216.4983	0.9197	224.0678	0.8497	247
8	259.0162	0.8437	269.4442	0.7092	283
9	285.8955	0.9086	298.3736	0.7906	319
10	326.2033	0.8613	342.1265	0.6787	355
11	366.0972	0.8122	385.8034	0.5647	391
12	444.5811	0.2689	472.4769	0.0634	427
13	482.4589	0.2570	514.6741	0.0484	463
14	506.1846	0.4023	541.3393	0.0926	499
15	543.2575	0.3930	583.3773	0.0726	535
16	560.3486	0.6169	602.9320	0.1717	571

\*The test is valid only for lags larger than the VAR lag order. df is degrees of freedom for (approximate) chi-square distribution

Pada Tabel 4.6.1. dapat dilihat bahwa hasil uji Portmanteau untuk *VECM(1)* pada *lag* 1, menyatakan bahwa tidak mengandung serial korelasi residual pada setiap *lag*, dimana semua nilai *Prob* dari statistik *Q* lebih dari taraf signifikansi  $\alpha = 5\%$  (artinya  $H_0$  diterima, dimana tidak ada serial korelasi). Ini artinya bahwa *VECM(1)* merupakan model terbaik.

## G. Uji Kausalitas

Berdasarkan bentuk persamaan model *VECM* di atas, diketahui bahwa variabel independen persamaan tersebut adalah IHSB. Pada persamaan tersebut dapat dilihat bahwa koefisien persamaan kointegrasi adalah 0.015631. Dengan kata lain, koefisien ECT bertanda positif. Berdasarkan hasil pengolahan data, diperoleh bahwa *pvalue* untuk statistik *t* adalah 0.04318 kurang dari taraf signifikansi  $\alpha = 5\%$ , yang berarti bahwa koefisien tersebut merupakan koefisien yang signifikan. Dengan demikian, koefisien ECT pada persamaan merupakan koefisien yang signifikan dan bertanda positif. Ini berarti bahwa, terdapat hubungan kausalitas jangka panjang (*long-run causality*) dari data tingkat Inflasi, Jumlah Uang Beredar, *BIRate*, Indeks Bursa Saham Global dan nilai tukar rupiah terhadap IHSB.

**Tabel 4.7.1** Uji Kausalitas Granger Variabel dependen IHSB

Dependent variable: D(IHSB)

Excluded	Chi-sq	df	Prob.
D(INFLASI)	0.922194	1	0.3369
D(BI_RATE)	0.934005	1	0.3338
D(KURS)	0.290751	1	0.5897
D(S_P)	0.726377	1	0.3941
D(M1)	1.402390	1	0.2363
All	5.228815	5	0.3886

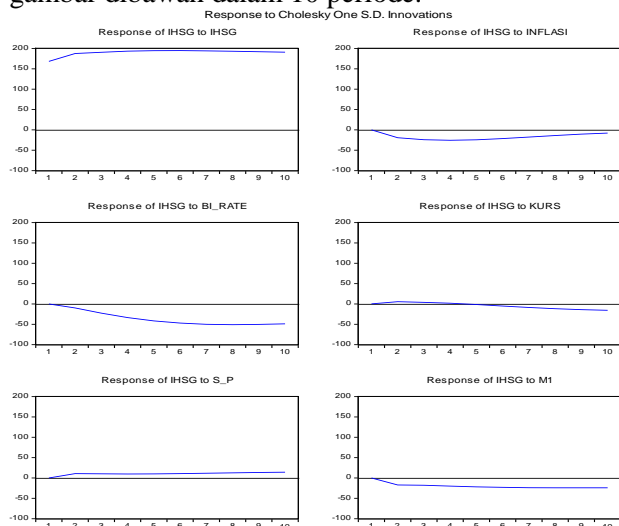
Untuk melihat hubungan kausalitas jangka pendek (*short-run causality*) dapat menggunakan Uji kausalitas Granger seperti yang diperlihatkan pada Tabel 4.7.1. Pada Tabel 4.7.1. dapat dilihat bahwa *p-value* statistik uji Wald untuk variabel D(INFLASI) adalah 0.3369 lebih dari taraf signifikan  $\alpha = 5\%$ . Ini berarti bahwa, menerima hipotesis  $H_0$  tidak ada hubungan kausalitas antara variabel. Dengan demikian, tidak terdapat hubungan kausalitas jangka pendek antara Inflasi terhadap IHSB. Sementara itu, untuk variabel D(BI\_rate) memiliki *p-value* adalah 0.3338 kurang dari taraf signifikan  $\alpha = 5\%$ . Ini berarti bahwa menerima hipotesis  $H_0$  pada taraf signifikansi  $\alpha = 5\%$  atau ada hubungan kausalitas jangka pendek antara BI Rate terhadap IHSB. Untuk variabel D(KURS) memiliki *p-value* adalah 0.5897 lebih dari taraf signifikan  $\alpha = 5\%$ . Ini berarti bahwa menerima hipotesis  $H_0$  pada taraf signifikansi  $\alpha = 5\%$  atau tidak ada hubungan kausalitas jangka pendek antara nilai tukar rupiah terhadap IHSB. Untuk variabel D(S\_P) memiliki *p-value* adalah 0.3941 lebih dari taraf signifikan  $\alpha = 5\%$ . Ini berarti bahwa menerima hipotesis  $H_0$  pada taraf signifikansi  $\alpha = 5\%$  atau tidak ada hubungan kausalitas jangka pendek antara Indeks Saham Global terhadap IHSB. Untuk variabel D(M1) memiliki *p-value* adalah 0.2363 lebih dari taraf signifikan  $\alpha = 5\%$ . Ini berarti bahwa menerima hipotesis  $H_0$  pada taraf signifikansi  $\alpha = 5\%$  atau tidak ada hubungan kausalitas jangka pendek antara jumlah uang beredar terhadap IHSB. Dan jika dilihat secara keseluruhan, maka memang tidak terdapat hubungan kausalitas jangka pendek dari tingkat Inflasi, Jumlah Uang Beredar, BI Rate, Indeks Bursa Saham Global dan nilai tukar rupiah terhadap IHSB. Hal ini dikarenakan oleh *p-value* = 0.3886 lebih besar dari taraf signifikan  $\alpha = 5\%$ , yang artinya bahwa menerima hipotesis  $H_0$ .

## H. IRF dan VD

### 1. Impulse Response Function

*Impulse Response Function* memberikan gambaran bagaimana respon dari suatu variabel dimasa mendatang jika terjadi gangguan pada satu

variabel lainnya. Untuk memudahkan interpretasi, hasil analisis disajikan dalam bentuk grafik pada gambar dibawah dalam 10 periode.



**Gambar 4.8.1. Impulse Response IHSB**

Jika grafik *impulse response* menunjukkan pergerakan yang semakin mendekati titik keseimbangan atau kembali ke keseimbangan sebelumnya, berarti respon suatu variabel akibat suatu guncangan makin lama makin menghilang sehingga guncangan tersebut tidak meninggalkan pengaruh permanen terhadap variabel tersebut.

Dari gambar di atas analisis *Impulse Response* dari tingkat IHSB untuk 10 periode ke depan sebagai berikut:

1. Respon IHSB terhadap guncangan diri sendiri pada periode 1 cukup besar, yaitu 1687,23. Meskipun hingga periode 6 sempat mengalami kenaikan menjadi 194,75, namun hingga kuartal 10 respon terhadap guncangan mengalami penurunan. Ini berarti setelah periode 6 reaksi IHSB terhadap guncangan diri sendiri cenderung menjadi lemah.
2. Respon IHSB terhadap guncangan tingkat inflasi bernilai negatif. Ini artinya jika terjadi guncangan 1 standar deviasi dari tingkat inflasi maka IHSB akan bereaksi negatif terhadap tingkat inflasi. Namun pada periode 7 mengalami peningkatan nilai, berarti respon IHSB terhadap guncangan tingkat inflasi cenderung meningkat dimulai dari periode 7 hingga periode 10.
3. Respon IHSB terhadap guncangan BI rate cenderung bernilai negatif. Ini artinya jika terjadi guncangan 1 standar deviasi dari BI rate maka IHSB akan bereaksi negatif terhadap BI rate.
4. Respon IHSB terhadap guncangan nilai tukar rupiah sempat meningkat pada periode 2, namun hingga periode 10 respon terhadap guncangan nilai tukar rupiah mengalami penurunan. Bahkan turun hingga bernilai negatif yang

berarti reaksi IHSB terhadap guncangan nilai tukar rupiah cenderung menjadi lemah.

5. Respon IHSB terhadap guncangan Indeks Global pada periode 1 hingga 10 mengalami peningkatan walaupun pada periode 4 sempat menurun sekali. Ini berarti bahwa guncangan sebesar 1 standar deviasi dari indeks global mengakibatkan kenaikan yang sangat signifikan terhadap IHSB.
6. Respon IHSB terhadap guncangan jumlah uang beredar (M1) cenderung bernilai negatif. Ini artinya jika terjadi guncangan 1 standar deviasi dari jumlah uang beredar maka IHSB akan bereaksi negatif terhadap jumlah uang beredar.

## 2. Variance Decomposition

*Variance decomposition* digunakan untuk menyusun *forecast error variance* suatu variabel, yaitu seberapa besar perbedaan antara varian sebelum dan sesudah guncangan, baik guncangan yang berasal dari diri sendiri maupun guncangan dari variabel lain untuk melihat pengaruh relatif variabel-variabel penelitian terhadap variabel lainnya. Prosedur *variance decomposition* yaitu dengan mengukur persentase kejutan-kejutan atas masing-masing variabel. Berikut ini disajikan *variance decomposition* untuk sepuluh periode kedepan atas masing-masing variabel pada tabel 4.8.1.

**Tabel 4.8.1. Variance Decomposition Variabel IHSB**

Variance Decomposition of IHSB:							
Period	S.E.	IHSB	INFLASI	BI_RATE	KURS	S_P	M1
1	168.2389	100.000	0.0000	0.0000	0.0000	0.0000	0.0000
2	253.8043	98.5863	0.5805	0.1492	0.0462	0.1791	0.4583
3	319.7963	97.6047	0.9343	0.6002	0.0442	0.2178	0.5985
4	376.7187	96.6703	1.1331	1.2264	0.0342	0.2237	0.7121
5	427.4232	95.8220	1.2055	1.9051	0.0277	0.2276	0.8119
6	473.2591	95.0949	1.1887	2.5465	0.0347	0.2360	0.8989
7	515.0521	94.5008	1.1219	3.0959	0.0585	0.2498	0.9728
8	553.4140	94.0357	1.0350	3.5307	0.0966	0.2682	1.0336
9	588.8539	93.6869	0.9461	3.8519	0.1431	0.2895	1.0821
10	621.8167	93.4374	0.8644	4.0739	0.1920	0.3121	1.1199

Tabel 4.8.1. merupakan rangkuman hasil analisis *Variance Decomposition* untuk IHSB dari guncangan yang diberikan oleh masing-masing variabel termasuk dirinya sendiri. Analisis *Variance Decomposition* pada tabel 4.8.1. menyatakan bahwa dalam jangka pendek, yaitu periode 3: guncangan terhadap dirinya sendiri mengakibatkan 97,60% fluktuasi dalam IHSB, guncangan terhadap indeks global mengakibatkan 0,21% fluktuasi dalam IHSB, guncangan terhadap jumlah uang beredar mengakibatkan 0,59% fluktuasi dalam IHSB, guncangan terhadap nilai tukar rupiah mengakibatkan 0,4% fluktuasi dalam IHSB,

guncangan terhadap tingkat inflasi mengakibatkan 0.93% fluktuasi dalam IHSG, guncangan terhadap BI rate mengakibatkan 0.60% fluktuasi dalam IHSG. Di lain pihak dalam jangka panjang, yaitu pada periode 10: guncangan terhadap dirinya sendiri mengakibatkan semakin lemah fluktuasi dalam IHSG, sedangkan guncangan terhadap indeks global, nilai kurs rupiah, jumlah uang beredar serta BI rate mengakibatkan fluktuasi semakin meningkat dalam IHSG. Selain itu juga guncangan inflasi mengakibatkan fluktuasi yang meningkat hingga periode 8 namun selanjutnya hingga periode 10 semakin melemah dalam IHSG.

### I. Peramalan

Setelah dilakukan pemodelan VECM dilanjutkan dengan tahap peramalan. Hasil ramalan terhadap IHSG untuk 3 periode kedepan dapat dilihat pada tabel 4.9.1 dan secara lengkap pada lampiran 10.

**Tabel 4.9.1.** Hasil Peramalan Untuk periode April 2016 hingga Juni 2016

Periode	Bulan	Ramalan IHSG
130	April 2016	4965.095
131	Mei 2016	4997.200
132	Juni 2016	5029.304

Untuk ukuran ketepatan ramalan dilakukan terhadap 129 data historis. Pengukuran kecocokan model menggunakan *Mean Absolute Error* (MAE) data indeks harga saham gabungan (IHSG) terhadap data tingkat inflasi, jumlah uang beredar, *BI Rate*, indeks bursa saham global, nilai tukar rupiah, dari bulan Juli 2005 s/d Maret 2016. Berdasarkan hasil perhitungan diperoleh bahwa nilai MAE dari data adalah 135.6353.

## V. PENUTUP

### A. Kesimpulan

Berdasarkan hasil dan pembahasan analisis peramalan menggunakan metode VECM, maka dapat disimpulkan bahwa:

1. Melalui uji kointegrasi dengan *Johansen's Cointegration Test* menunjukkan bahwa seluruh variabel penelitian yaitu data tingkat inflasi, jumlah uang beredar, *BI Rate*, indeks bursa saham global, nilai tukar rupiah, dan indeks harga saham gabungan (IHSG) dalam setiap periode jangka pendek cenderung saling menyesuaikan untuk mencapai ekuilibrium jangka panjangnya.
2. Berdasarkan spesifikasi model (analisis *lag* optimal) maka diperoleh model terbaik untuk data adalah VECM (1), dengan model VECM(1) sebagai berikut:

$$\begin{aligned}
 D(IHSG) = & 0.016 * (IHSG(-1) - 18041.633 \\
 & * INFLASI(-1) + 21491.578 \\
 & * BI\_RATE(-1) + 0.373 \\
 & * KURS(-1) - 1.554 * S\_P(-1) \\
 & + 0.002 * M1(-1) - 50.076 \\
 & * @TREND(05M07) \\
 & - 3176.334) + 0.058 \\
 & * D(IHSG(-1)) - 1253.786 \\
 & * D(INFLASI(-1)) - 6295.125 \\
 & * D(BI\_RATE(-1)) + 0.029 \\
 & * D(KURS(-1)) + 0.285 \\
 & * D(S\_P(-1)) - 0.0009 \\
 & * D(M1(-1)) + 30.409
 \end{aligned}$$

Interpretasi persamaan diatas untuk nilai IHSG pada saat t dipengaruhi oleh enam variabel, dimana tiap variabel memberikan kontribusi yang berbeda. Beberapa variabel yang mempengaruhi secara positif adalah variabel IHSG, Kurs, dan Indeks Saham Global (S&P 500). Sedangkan, variabel yang mempengaruhi secara negatif adalah variabel Inflasi, *BI Rate* dan Jumlah Uang Beredar. Berdasarkan hal tersebut untuk meningkatkan poin IHSG pada periode berikutnya agar meningkatkan hasil pada variabel yang bersifat positif dan mengendalikan variabel yang bersifat negatif pada bulan sebelumnya (*lag(-1)*).

3. Berdasarkan analisis struktural dari VECM(1) yaitu IRF dan VD, dapat disimpulkan bahwa:
  - a. Respon IHSG terhadap guncangan diri sendiri dan indeks saham global cukup signifikan, karena terjadi fluktuasi.
  - b. Respon dari IHSG terhadap guncangan dari jumlah uang beredar, nilai tukar rupiah, dan *BI rate* tidak signifikan.
4. Hasil ramalan yang diperoleh dengan menggunakan VECM(1) menghasilkan nilai ramalan pada bulan April 2016 hingga Juni 2016 berturut-turut sebesar 4965.095, 4997.200 dan 5029.304. Berarti untuk 3 periode ke depan IHSG mengalami peningkatan nilai.

### B. Saran

Dari hasil pembahasan dan kesimpulan yang didapat, maka saran yang Peneliti sampaikan adalah:

1. Peneliti menyarankan untuk mengkaji lebih dalam mengenai hubungan antara IHSG dengan faktor-faktor yang mempengaruhinya, baik secara teori maupun menggunakan metode.
2. Dalam analisis terdapat permasalahan multikolinieritas yang mengakibatkan adanya near singular matrix saat melakukan eksekusi pada program *eviews*, sehingga ada baiknya untuk melihat terlebih dahulu apakah variabel bebas memiliki hubungan yang kuat yang dapat mengakibatkan timbulnya multikolinieritas.

## DAFTAR PUSTAKA

- Adu, I. S. 2013. Error Correction Model: Dampak Kebijakan Fiskal, Kebijakan Moneter, Kestabilan Perekonomian Nasional serta Dunia terhadap Pertumbuhan Ekonomi Indonesia. *Jurnal Makroekonomi, Sekolah Tinggi Ilmu Statistik (STIS)*.
- Alwi, I. 2003. *Pasar Modal : Teori dan Aplikasi*. Jakarta: Nasindo Internusa.
- Amrulloh, M. A. 2013. *Pengaruh Inflasi, Suku Bunga Sbi, Nilai Tukar Rupiah, Harga Minyak Dunia Serta Indeks Dow Jones Terhadap Indeks Harga Saham Gabungan Di Bursa Efek Indonesia Periode 2010-2013*. Skripsi. Universitas Negeri Malang.
- Anonim. 2008. *Analisis Hubungan Kointegrasi dan Kausalitas serta Hubungan Dinamis antara Aliran Modal Asing, Perubahan Nilai Tukar dan Pergerakan IHSG di Pasar Modal Indonesia*. Jakarta: Badan Pengawas Pasar Modal dan Lembaga Keuangan Departemen Keuangan Republik Indonesia.
- Anonim. 2010. *Ekuitas*. Website Bursa Efek Indonesia. Diakses tanggal 25 Juli 2015. <http://www.idx.co.id/id-id/beranda/informasi/bagiinvestor/saham.aspx>
- Anonim. 2010. *Pengantar Pasar Modal*. Website Bursa Efek Indonesia. Diakses tanggal 25 Juli 2015. <http://www.idx.co.id/id-id/beranda/informasi/bagiinvestor/pengantarpasarmodal.aspx>
- Anonim. 2013. *EViews 8 User's Guide II*. ISBN: 978-1-880411-13-1. IHS Global Inc.
- Box, G.E.P and G.M. Jenkins. 1994. *Time Series Analysis: Forecasting and Control*. Third Edition. Prentice Hall: New Jersey.
- Brooks, C. 2008. *Introductory Econometrics for Finance 2nd Edition*. New York: Cambridge University Press.
- Divianto. 2013. Analisis Pergerakan Tingkat Inflasi, Tingkat Suku Bunga SBI, dan Nilai Kurs Dollar AS (USD) terhadap Indeks Harga Saham Gabungan (IHSG) di Bursa Efek Indonesia. *Jurnal Ekonomi dan Informasi Akutansi (JENIUS)*, 3(2).
- Enders, W. 2004. *Applied Econometric Time Series Second Edition*. Hoboken: John Wiley and Son, Inc.
- Engle, R.F. dan C. W. J. Granger. 1987. Co-integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, Vol. 55, No. 2, March 251-276.
- Gujarati, D. 2004. *Basic Econometrics, Fourth Edition*. New York: McGraw-Hill.
- Halim, S., & Chandra, A. 2011. Pemodelan Time Series Multivariat secara Automatis. *Jurnal Teknik Industri*, 13(1), 19-26.
- Indrawati, F dan B. Sutijo. 2012. "Pemodelan Jumlah Ketersediaan Beras untuk Jawa Timur Dengan Pendekatan Fungsi Transfer". *Jurnal Sains dan Seni ITS* Vol. 1, No. 1.
- Lütkepohl, H. 2005. *New Introduction to Multiple Time Series Analysis*. Springer-Verlag. Berlin.
- Makridakis, S., S.C Wheelwright dan V. McGee. Alih bahasa Ir. Hari Suminto. 1999. *Metode dan Aplikasi Peramalan*. Edisirevisi. Binarupa Aksara: Jakarta.
- Prakoso, A. T. 2009. *Analisis Hubungan Perdagangan Internasional dan FDI terhadap Pertumbuhan Ekonomi Indonesia*. Depok: Fakultas Ekonomi, Universitas Indonesia.
- Purba, B. 2014. Analisis Kointegrasi Antara Indeks Harga Saham Gabungan (IHSG), Jumlah Uang Beredar (JUB), dan Indeks Harga Pedagang Besar (IHPB) di Indonesia Periode Tahun 007-2013. *Jurnal Saintecth*, Vol. 06, No 4.
- Sari, S.P. 2015. *Peramalan Nilai Inflasi terhadap Nilai Tukar Rupiah Indonesia dengan Pendekatan VAR*. Skripsi. Universitas Bengkulu.
- Sims, C.A. 1980. Macroeconomics and Reality. *Econometrica*, Vol. 48, No. 1 (Jan., 1980), 1-48.
- Subagyo, A. 2013. Analisis ECM dan VECM pada Variabel Makroekonomi di Indonesia. *Wawasan Tridharma: Majalah Ilmiah Kopertis IV Nomor 01 Tahun XXVI*, Jakarta.
- Syahmer, V. 2010. *Keterkaitan Nilai Tukar Rupiah Dengan Indeks Saham di Bursa Efek Indonesia*. MB-IPB.
- Wei, W.W.S. 2006. *Time Series Analysis: Univariate and Multivariate Methods*. Addison-Wesley Publishing Company: New York.
- Wijaya, T. 2013. *Pengaruh Berbagai Faktor Internal dan Eksternal terhadap Pergerakan Indeks Harga Saham Gabungan (IHSG) di Bursa Efek Indonesia*. Palembang: Fakultas Ekonomi, Universitas Sriwijaya.

# ANALISIS JALUR FAKTOR-FAKTOR PENYEBAB KRIMINALITAS DI INDONESIA

Ria Pasiza<sup>1</sup>, Sigit Nugroho<sup>2</sup>, dan Fachri Faisal<sup>2</sup>

Mahasiswa Jurusan Statistika, Prodi Matematika, FMIPA, UNIB, Bengkulu  
Dosen Pembimbing Jurusan Statistika, Prodi Matematika, FMIPA, UNIB, Bengkulu  
[riapasizza@gmail.com](mailto:riapasizza@gmail.com)

## ABSTRAK

Penelitian ini bertujuan untuk mengetahui pengaruh faktor-faktor penyebab kriminal di Indonesia, dengan membuat paradigma penelitian, menguji koefisien jalur dan mencari korelasinya, mengetahui hubungan kausal antar variabel eksogen dan endogen terhadap kriminalitas. Penelitian merupakan studi kasus dengan menggunakan data sekunder dari publikasi Badan Pusat Statistik provinsi Bengkulu yang meliputi data Kependudukan, Kemiskinan, Ketenagakerjaan, dan Pidana. Hasil penelitian menunjukkan bahwa variabel yang berpengaruh secara langsung terhadap kriminalitas di Indonesia adalah persentase kepadatan penduduk, tingkat pengangguran terbuka, dan persentase penduduk miskin. Pengaruh secara tidak langsung terhadap kriminalitas di Indonesia adalah persentase kepadatan penduduk tingkat pengangguran terbuka.

Kata kunci : *Analisis Jalur, Koefisien Jalur, kriminalitas, Variabel Eksogen, Variabel Endogen.*

## PENDAHULUAN

Indonesia adalah salah satu negara yang mempunyai jumlah penduduk terbesar di dunia. Jumlah penduduk Indonesia adalah 255.609.643 juta jiwa. Berdasarkan jumlah penduduk tersebut, Indonesia menempati urutan ke-4 dunia setelah Amerika Serikat, India, dan Cina. Padatnya jumlah penduduk mengakibatkan terjadinya permasalahan dalam kehidupan masyarakat Indonesia dan merupakan masalah yang sangat kompleks bagi pemerintahan pada saat ini (Anonim, 2014).

Kriminalitas merupakan salah satu permasalahan yang dihadapi bangsa Indonesia, kesuksesan pembangunan yang diusahakan pemerintah oleh setiap negara sangat bergantung terhadap besar kecilnya hambatan dari kriminalitas, baik pembangunan sumber daya manusia itu sendiri maupun infrastruktur. Peran aktif dan dukungan dari masyarakat sangat dibutuhkan terhadap proses laju pembangunan akan maksimal jika tingkat kriminalitas berkurang. Berbagai kerugian telah banyak ditimbulkan oleh adanya tindak kriminal, baik itu kerugian ekonomi, fisik, moral, dan psikologis. Dari sudut pandang ekonomi, kriminalitas menimbulkan kerugian dengan adanya biaya yang harus dikeluarkan akibat dari tindak kriminal. Biaya tersebut tidak hanya ditanggung korban, namun oleh masyarakat, dunia usaha, dan juga negara atau pemerintah daerah (Sullivan, 2009).

Penelitian sebelumnya tentang tingkat kriminalitas pernah dilakukan oleh Hakim (2009) yang melakukan analisis determinan terhadap tingkat kejahatan di Jawa Tahun 2007. Selain itu terdapat pula Hardianto (2009) yang menganalisis faktor-faktor yang mempengaruhi tingkat kriminalitas di Indonesia dari pendekatan ekonomi. Putra (2008) juga melakukan identifikasi faktor-faktor yang mempengaruhi tindak kriminal berdasarkan karakteristik pelaku kriminal dengan metode pohon klasifikasi. Menurut Masliyah (2012) kriminalitas umumnya dipengaruhi oleh faktor internal dan eksternal yang secara langsung mempengaruhi tingkat kriminalitas dengan menggunakan pendekatan simulasi sistem dinamik. Faktor-faktor penting tersebut adalah tingkat pendidikan yang diperoleh oleh pelaku kriminalitas, kesejahteraan berpengaruh pada terpenuhinya atau tidaknya kebutuhan material dan pendapatan yang diperoleh. Pengaruh lingkungan seperti keluarga dan teman sebaya juga menjadi salah satu faktor timbulnya tindak kriminalitas dan faktor penyebab kriminalitas yang terakhir adalah pengangguran (*unemployment*).

Pada penelitian juga telah dilakukan oleh Mariana dan Budiantara (2013) tentang faktor-faktor yang mempengaruhi persentase kriminalitas di Jawa Timur yaitu kepadatan penduduk, tingkat pengangguran terbuka, persentase penduduk miskin, persentase penduduk yang tidak pernah sekolah, persentase penduduk yang merupakan korban penyalahgunaan NAPZA, dan persentase keluarga

bermasalah dengan menggunakan pendekatan regresi semiparametrik spline. Metode analisis jalur juga pernah dilakukan oleh (Zhang, 2014) dalam menguji hubungan antara variabel pembelajaran *multitasking* dan prestasi akademik pada beberapa mahasiswa di Singapura, dalam penelitian analisis jalur digunakan untuk menunjukkan bahwa perilaku di kelas *multitasking* memberikan pengaruh yang negatif terhadap pendidikan siswa. (Kurniati, 2011) menggunakan analisis jalur untuk mengetahui faktor-faktor yang mempengaruhi tingkat pengangguran di Indonesia.

Salah satu analisis statistika yang dapat digunakan untuk menganalisa masalah ini adalah analisis jalur. Analisis jalur merupakan bagian dari analisis statistika multivariat yang digunakan untuk menganalisis pola hubungan antar variabel dengan tujuan untuk mengetahui pengaruh langsung dan tidak langsung seperangkat variabel bebas (*eksogen*) terhadap variabel terikat (*endogen*). Sehingga dalam penelitian ini dapat dijelaskan variabel-variabel yang berpengaruh langsung dan tidak langsung penyebab kriminalitas di Indonesia.

## Tinjauan Pustaka

### Analisis Jalur (Path Analysis)

Analisis jalur yang dikenal dengan *Path Analysis* dikembangkan pertama pada tahun 1920-an oleh seorang ahli genetika yaitu Sewall Wright (Riduwan dan Engkos, 2012). Teknik analisis jalur merupakan perkembangan korelasi yang diuraikan menjadi beberapa interpretasi akibat yang ditimbulkannya. Analisis jalur mempunyai kedekatan dengan regresi berganda, yang merupakan bentuk khusus dari analisis jalur. Teknik ini dikenal sebagai model sebab akibat (Sarwono, 2007). Analisis jalur dapat digunakan untuk menguji model teoritis yang menentukan hubungan antara arah sejumlah variabel yang diamati. Analisis jalur menentukan model hubungan yang sebenarnya dari data sampel yang diamati (Zhang, 2014).

### Asumsi-Asumsi Analisis Jalur

sebelum melakukan analisis, maka perlu diperhatikan beberapa asumsi yang mendasari analisis jalur di bawah ini :

1. Hubungan antar variabel bebas dan tak bebas di dalam model adalah linier artinya perubahan yang terjadi pada variabel adalah merupakan fungsi perubahan linier dari variabel lainnya yang bersifat kausal.
2. Variabel yang diamati mempunyai sifat aditif artinya variabel yang mempunyai sifat multiplikatif dan eksponensial tidak dapat dipergunakan.
3. Variabel sisa tidak berkorelasi dengan variabel yang sesudahnya (variabel regresi lainnya).
4. Variabel yang diukur berskala interval atau rasio.

### Koefisien Jalur

Menurut Dillon & Goldstein (1984) Koefisien jalur mengindikasikan besarnya pengaruh langsung dari suatu variabel yang mempengaruhi terhadap variabel yang dipengaruhi atau dari suatu variabel eksogen terhadap variabel endogen. Simbol atau notasi konvensional untuk melambangkan koefisien jalur adalah  $\rho_{ij}$ , dimana  $i$  merefleksikan akibat (*dependent variable*) dan merefleksikan sebab (*independent variable*).

Menurut Wibowo (2005) menggambarkan konsep koefisien jalur sebagai berikut. Misalkan ada  $m$  variabel bebas, yakni  $X_1, X_2, \dots, X_m$  dan  $Y$  sebagai variabel tak bebas.  $\xi$  merupakan notasi untuk variabel sisa (*residual*) dan semua variabel saling berkorelasi kecuali variabel sisa. Koefisien  $C_{y x_i}$  ( $i = 1, 2, \dots, m$ ) menunjukkan sumbangan nyata  $X_i$  secara langsung terhadap  $Y$  serta semua hubungan adalah linier, maka persamaannya adalah

$$Y = C_{y x_1} X_1 + C_{y x_2} X_2 + \dots + C_{y x_m} X_m + C_y \xi \quad (1)$$

Jika  $X_i = \frac{\sigma_i}{\sigma_0} V_i, \forall i, i = 0, 1, 2, \dots, m$

Keterangan :

$V_i$  = variabel ke- $i$  ( $i = 0, 1, 2, \dots, m$ )

$\sigma_i$  = simpangan populasi dari variabel bebas ke- $i$

$\sigma_y$  = simpangan baku populasi dari variabel tak bebas

Maka,

$$Y = C_{y x_1} \left( \frac{\sigma_1}{\sigma_y} \right) V_1 + C_{y x_2} \left( \frac{\sigma_2}{\sigma_y} \right) V_2 + \dots + C_{y x_m} \left( \frac{\sigma_m}{\sigma_y} \right) V_m + C_y \left( \frac{\sigma_\xi}{\sigma_y} \right) \xi \quad (2)$$

Apabila  $\gamma_{y x_i} = C_{y x_i} \left( \frac{\sigma_i}{\sigma_y} \right)$  maka,

$$Y = \gamma_{y x_1} V_1 + \gamma_{y x_2} V_2 + \dots + \gamma_{y x_m} V_m + \varepsilon \quad (3)$$

$\gamma$  dengan subskrip ganda menunjukkan hubungan antara variabel tak bebas dan variabel bebas. Oleh karena itu  $\gamma$  adalah koefisien jalurnya dan merupakan koefisien yang dipergunakan untuk menghitung arah panah yang tersusun dalam model (Nugroho, 2008).

#### Koefisien Determinasi dan Koefisien Residu

Koefisien determinasi  $R^2$  adalah besarnya pengaruh bersama-sama variabel eksogen terhadap variabel endogen yang dapat dijelaskan oleh model persamaan jalur. Nilai  $R^2$  persamaan jalur yang makin mendekati 100% menunjukkan bahwa makin banyak keragaman variabel eksogen terhadap variabel endogen yang dapat dijelaskan dari persamaan jalur tersebut (Rahayu, 2013).

Dalam bentuk baku, semua koefisien korelasi direduksi dari hasil kali momen yaitu (Nugroho, 2008):

$$\begin{aligned} r_{y x} &= \frac{\sum_{i=1}^m Y V_i}{n} \\ &= \frac{1}{n} \left( \sum_{i=1}^m (\gamma_{y x_1} V_1 + \gamma_{y x_2} V_2 + \dots + \gamma_{y x_m} V_m + \gamma_{y x_u} V_u) V_i \right) \\ &= \frac{1}{n} \left( \sum_{i=1}^m (\gamma_{y x_1} V_1 V_i + \gamma_{y x_2} V_2 V_i + \dots + \gamma_{y x_m} V_m V_i + \gamma_{y x_u} V_u V_i) \right) \\ &= \gamma_{y x_1} \sum_{i=1}^m \frac{V_1 V_i}{n} + \gamma_{y x_2} \sum_{i=1}^m \frac{V_2 V_i}{n} + \dots + \gamma_{y x_m} \sum_{i=1}^m \frac{V_m V_i}{n} + \gamma_{y x_u} \sum_{i=1}^m \frac{V_u V_i}{n} \\ &= \gamma_{y x_1} r_{x_1 x_i} + \gamma_{y x_2} r_{x_2 x_i} + \dots + \gamma_{y x_m} r_{x_m x_i} + \gamma_{y x_u} r_{x_u x_i} \\ &= \sum_{i=1}^m \sum_{j=1}^m \gamma_{y x_j} r_{x_j x_i} \quad (4) \end{aligned}$$

Koefisien determinasi dapat ditulis dalam bentuk matriks sebagai berikut :

$$R^2 = (\gamma_{y x_1} \gamma_{y x_2} \dots \gamma_{y x_k}) \begin{pmatrix} r_{y x_1} \\ r_{y x_2} \\ \vdots \\ r_{y x_k} \end{pmatrix} \quad (5)$$

Berdasarkan koefisien determinasi dapat diidentifikasi pengaruh koefisien residual atau faktor sisaan. Koefisien residu  $\varepsilon$  adalah besarnya pengaruh variabel lain diluar model yang tidak ikut diamati.

Dari persamaan (6) , bila  $i = 0$  maka,

$$r_{0 0} = \gamma_{y x_1} r_{x_1 x_0} + \gamma_{y x_2} r_{x_2 x_0} + \dots + \gamma_{y x_m} r_{x_m x_0} + \gamma_{y x_u} r_{x_u x_0}$$

jika  $\gamma_{y x_u} = r_{x_u x_0}$  dan karena  $r_{0 0} = 1$  maka

$$1 = \gamma_{y x_1} r_{x_1 x_0} + \gamma_{y x_2} r_{x_2 x_0} + \dots + \gamma_{y x_m} r_{x_m x_0} + \gamma_{y x_u}^2$$

$$\gamma_{y x_u}^2 = 1 - (\gamma_{y x_1} r_{x_1 x_0} + \gamma_{y x_2} r_{x_2 x_0} + \dots + \gamma_{y x_m} r_{x_m x_0})$$

$$\gamma_{y x_u}^2 = 1 - \sum_{i=1}^m \gamma_{y x_i} r_{x_i x_0}$$

$$\gamma_{y x_u} = \sqrt{1 - \sum_{i=1}^m \gamma_{y x_i} r_{x_i x_0}} \quad (6)$$

Persamaan (8) merupakan rumus untuk residual atau galat untuk setiap variabel endogen.

#### Dekomposisi Korelasi

Suatu penguraian korelasi antar variabel terhadap variabel endogen disebut dengan dekomposisi korelasi. Dekomposisi ini adalah perbedaan mendasar analisis jalur dengan analisis regresi berganda (Nugroho, 2008).

Menurut Gaspersz (1992) apabila koefisien jalur telah diperoleh berdasarkan metode analisis jalur, antara lain :

1. Pengaruh langsung (*Direct Effect/ DE*) variabel bebas terhadap variabel tidak bebas.
2. Pengaruh tidak langsung (*Indirect Effect/IE*) variabel bebas terhadap variabel tidak bebas.

- Pengaruh galat (Error) atau sisaan (residual), yang tidak dapat dijelaskan oleh model analisis jalur (pengaruh-pengaruh yang tidak dapat dijelaskan oleh suatu model dimasukkan sebagai galat atau sisaan).

### Uji Kesesuaian Model (*Overall Model Fit*)

Pengujian kecocokan model (*model fit*) dapat digunakan statistik *Chi-Square* yang dianjurkan oleh Pedhazur (1982). Suatu model yang diusulkan dikatakan cocok dengan data bila matriks korelasi model teoritis sama dengan matriks korelasi empiris (*Reproduced*). Dengan demikian, perumusan hipotesis pada analisis jalur ditulis sebagai berikut :

$$H_0 : R = R(\theta)$$

$$H_1 : R \neq R(\theta)$$

Model dikatakan cocok atau *fit* jika hipotesis nol diterima. Untuk menguji hipotesis tersebut dapat digunakan statistic *Chi-Square* yaitu :

$$W = -(n - d) \ln(Q) \quad (7)$$

Dimana  $n$  menunjukkan ukuran sampel dan  $d$  menunjukkan banyaknya koefisien jalur yang sama dengan nol atau koefisien jalur yang *nonsignificant*, dan  $Q$  adalah :

$$Q = \frac{1 - R_m^2}{1 - M} \quad (8)$$

Dimana  $R_m^2$  adalah koefisien determinasi multiple untuk model yang diusulkan dan  $M$  adalah koefisien determinasi multiple untuk model setelah terdapat koefisien jalur yang *nonsignificant*. Koefisien determinasi tersebut adalah :

$$M = R_m^2 = 1 - (1 - R_1^2)(1 - R_2^2) \cdots (1 - R_p^2) \quad (9)$$

Statistik  $W$  mendekati distribusi *Chi-Square* dengan derajat bebas sebesar  $d$ . Jika nilai  $W$  sangat kecil atau mendekati nilai nol, maka hipotesis nol diterima. Dengan kata lain bahwa model yang diusulkan cocok dengan data.

Selain menggunakan statistic *Chi-Square* pengujian model juga dapat dilakukan dengan menggunakan indeks kesesuaian model. Berikut ini beberapa indeks kesesuaian model dan *cut of value*-nya yang dapat digunakan untuk menguji sebuah model dapat diterima atau ditolak (Nugroho, 2008):

- RMSEA (*The Root Mean Square Error of Approximation*)

RMSEA adalah sebuah indeks yang dapat digunakan untuk mengkompensasi *Chi-Square Statistics* dalam sampel yang besar (Baumgartner & Homburg, 1996). Nilai RMSEA yang lebih kecil atau sama dengan 0,08 merupakan indeks untuk dapat diterimanya model yang menunjukkan sebuah *close fit* dari model berdasarkan *degrees of freed* (Browne & Cudeck, 1993).

- GFI (*Goodness of Fit Index*)

GFI adalah sebuah ukuran non-statistical yang mempunyai rentang nilai antara 0 (*poor fit*) sampai dengan 1,0 (*perfect fit*). Nilai yang tinggi dalam indeks ini menunjukkan sebuah '*better fit*'.

- AGFI (*Adjusted Goodness of Fit Index*)

Tanaka & Huba (1989) menyatakan bahwa GFI adalah analog dengan  $R^2$  dalam regresi berganda. Nilai sebesar 0,95 dapat diinterpretasikan sebagai tingkatan yang baik (*good overall model fit*) sedangkan besaran nilai antara 0.90 – 0.95 menunjukkan tingkatan cukup (*adequate fit*).

- CMIN (*The Minimum Sample Discrepancy Function*)/DF (*Degree of Freedom*)

CMIN dibagi dengan *Degree of Freedom*-nya akan menghasilkan indeks CMIN/DF, yang umumnya dilaporkan oleh para peneliti sebagai salah satu indikator untuk mengukur tingkat *fit*-nya sebuah model. Dalam hal ini CMIN/DF adalah statistik *Chi-Square*,  $\chi^2$  dibagi DF-nya sendiri sehingga disebut  $\chi^2$ -relatif. Nilai  $\chi^2$ -relatif kurang dari 2,0 atau bahkan kurang dari 3,0 adalah indikasi dari *acceptable fit* antara model dan data.

- TLI (*Tucker Lewis Index*)

TLI adalah sebuah alternatif *incremental fit index* yang membandingkan sebuah model yang diuji terhadap sebuah *baseline model* (Baumgartner & Homburg, 1996). Nilai yang direkomendasikan sebagai acuan untuk diterimanya sebuah model adalah  $\geq 0,95$  (Hair dkk, 1995), dan nilai yang sangat mendekati 1 menunjukkan a *very good fit* (Arbuckle, 1997).

- CFI (*Comparative Fit Index*)

Besaran indeks ini adalah pada rentang nilai sebesar 0-1. Semakin mendekati mengindikasikan tingkat *fit* yang paling tinggi (Arbuckle, 1997). Sedangkan nilai yang

direkomendasikan adalah  $CFI \geq 0,95$ . Keunggulan dari indeks ini adalah bahwa indeks ini besarnya tidak dipengaruhi oleh ukuran sampel karena ia sangat baik untuk mengukur tingkat penerimaan sebuah model (Hulland, Chow dan Lam, 1996).

### Metode Estimasi

Estimasi adalah menaksir ciri-ciri tertentu dari populasi atau memperkirakan nilai populasi (parameter) dengan memakai nilai sampel (statistik). Untuk mendekati kesesuaian (*Fit*) antara model dengan data maka piranti lunak berbasis SEM akan mengestimasi nilai semua parameter. *AMOS* akan melakukan estimasi *koefisien path*. Dalam melakukan estimasi model, ukuran sampel memegang peranan yang cukup penting.

Dalam program *AMOS* teknik-teknik estimasi yang tersedia adalah :

- a) *Maximum Likelihood* (ML)
- b) *Generalized Least Square* (GLS)
- c) *Unweighted Least Square* (ULS)
- d) *Scale Free Least Square* (SLS)
- e) *Asymtotically Distribution-free Estimation* (ADF)

Suatu cara yang penting untuk mendapat estimator yang baik adalah metode maksimum likelihood yang diperkenalkan oleh R. A. Fisher. Maksimum likelihood merupakan salah satu pendekatan terpenting pada penaksiran dalam semua statistik inferensia. Metode estimasi lainnya misalkan *Generalized Least Square Estimation* (GLS) sering digunakan apabila pertimbangan asumsi normalitas dipenuhi dan ukuran sampel antara 200-500 dan bila ukuran sampel kurang dari 500. *Generalized Least Square Estimation* (GLS) lebih sering menerima model yang salah (*false*) daripada ML, metode ini digunakan sebagai konsekuensi dari ketidakakuratan maka indeks modifikasi (MI). *general least square* merupakan estimator yang tahan terhadap data tidak normal, maka pada penelitian ini digunakan metode estimasi *Generalized Least Square Estimation* (GLS).

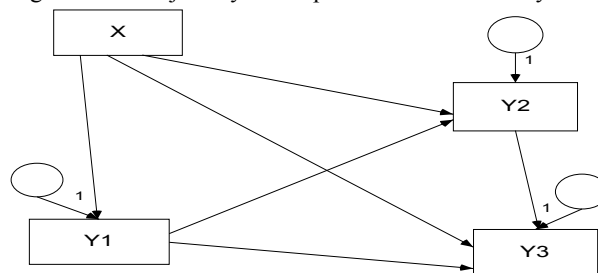
## METODE PENELITIAN

### Data

Penelitian ini menggunakan data sekunder yang diambil dari publikasi BPS (Badan Pusat Statistik) Bengkulu. Penulis memperoleh data yang tersedia di perpustakaan BPS provinsi Bengkulu yang merupakan data dari 31 provinsi di buku Statistik Indonesia tahun 2013, sehingga terdapat 31 sampel yang mewakili masing-masing provinsi. Variabel yang diamati pada penelitian ini adalah  $X$ : persentase kepadatan penduduk (%),  $Y_1$ : tingkat pengangguran terbuka (%),  $Y_2$ : persentase penduduk miskin (%),  $Y_3$ : jumlah kriminal (ribu).

### Analisis Data

Data yang diperoleh akan dianalisis dengan menggunakan analisis jalur (path analysis) dengan menggunakan software *Amos 16*. Berdasarkan paradigma hubungan antar variabel maka model diagram jalur yang dilengkapi dengan koefisien jalurnya serta persamaan strukturalnya.



Persamaan strukturalnya adalah sebagai berikut :

$$\begin{aligned}
 Y_1 &= \gamma_1 X + \varepsilon \\
 Y_2 &= \gamma_2 X + \beta_1 Y_1 + \varepsilon \\
 Y_3 &= \gamma_3 X + \beta_2 Y_1 + \beta_3 Y_2 + \varepsilon
 \end{aligned}$$

## Hasil dan Pembahasan

Analisis diawali dengan pengujian asumsi normalitas dan linieritas terhadap data dengan menggunakan *Amos 16* dan hasil analisis data tidak normal, maka peneliti menggunakan metode estimasi GLS (*Generalized Least Square*).

### a. Perhitungan Koefisien Pengaruh Langsung antar variable

Pengaruh langsung (*Dirrect Effect*) adalah efek atau pengaruh satu variabel ke variabel lain tanpa melalui satu variabel atau lebih. Pengaruh tidak langsung (*Indirect Effect*) adalah pengaruh yang muncul melalui sebuah variabel antara. Sedangkan efek total (*Total Efect*) adalah pengaruh dari berbagai hubungan.

Tabel 1. Standardized Direct Effects

	x	y1	y2
y1	,201	,000	,000
y2	-,245	,170	,000
y3	,162	,251	-,041

Dari tabel di atas dapat dilihat bahwa nilai standardized pengaruh langsung (*Direct Effect*) pengaruh  $X$  (persentase kepadatan penduduk) terhadap  $Y_1$  (tingkat pengangguran terbuka) adalah sebesar 0,201,  $X$  (persentase kepadatan penduduk) terhadap  $Y_2$  (persentase penduduk miskin) adalah  $-0,245$ ,  $X$  (persentase kepadatan penduduk) terhadap  $Y_3$  (jumlah kriminal) adalah 0,162,  $Y_1$  (tingkat pengangguran terbuka) terhadap  $Y_2$  (persentase penduduk miskin) adalah 0,170,  $Y_1$  (tingkat pengangguran terbuka) terhadap  $Y_3$  (jumlah kriminal) adalah 0,251,  $Y_2$  (persentase penduduk miskin) terhadap  $Y_3$  (jumlah kriminal) adalah  $-0,041$ .

**b. Perhitungan Koefisien Pengaruh Tidak Langsung antar variable**

Pada dasarnya mengestimasi koefisien yang mempengaruhi dalam konteks analisis jalur dan bukan hanya menganalisis pengaruh langsung (*Direct Causal Effect*), tetapi juga sekaligus mengestimasi pengaruh tidak langsung (*Indirect Causal Effect*), dan pengaruh total (*Total Causal Effect*).

Tabel 2. Standardized Indirect Effects

	x	y1	y2
y1	,000	,000	,000
y2	,034	,000	,000
y3	,059	-,007	,000

Dari tabel 2 dapat dilihat pengaruh tidak langsung  $X$  (persentase kepadatan penduduk) terhadap  $Y_2$  (persentase penduduk miskin) melalui  $Y_1$  (tingkat pengangguran terbuka) adalah sebesar 0,034 dengan rincian path tidak langsung  $X$  terhadap  $Y_2$  melalui  $Y_1$  (tingkat pengangguran terbuka) yaitu dari  $X$  ke  $Y_1$  (0,201) dan dari  $Y_1$  ke  $Y_2$  (0,170). Dengan demikian pengaruh tidak langsung  $X$  (persentase kepadatan penduduk) terhadap  $Y_2$  (persentase penduduk miskin) adalah  $= (0,201) \times (0,170) = 0,034$ . Pengaruh tidak langsung  $X$  (persentase kepadatan penduduk) terhadap  $Y_3$  melalui  $Y_1$  (tingkat pengangguran terbuka) yaitu sebesar (0,0504) dan dari  $X$  (persentase kepadatan penduduk) menuju  $Y_3$  melalui  $Y_2$  (persentase penduduk miskin) adalah sebesar  $(-0,0014)$ , dari  $X$  menuju  $Y_3$  melalui  $Y_2$  adalah sebesar 0,0100 dengan menjumlahkan nilai dari masing-masing jalur yang dilalui maka didapatkan besar pengaruh tidak langsung jalur  $X$  (persentase kepadatan penduduk) menuju  $Y_3$   $= (-0,0014 + 0,0504 + 0,0100) = 0,059$ . Pengaruh tidak langsung  $Y_1$  (tingkat pengangguran terbuka) terhadap  $Y_3$  melalui  $Y_2$  (persentase penduduk miskin) adalah sebesar  $= (0,170) \times (-0,041) = -0,07$ .

**c. Pengujian Model dengan Kriteria Kesesuaian**

Dari hasil output program amos diperoleh nilai *Chi-Square* sebesar *Chi-Square* = 0.00 dan  $P$  adalah  $1(P > 0.05)$ . Menurut Ghazali dan Fuad (2008), nilai *Chi-Square* sebesar 0 menunjukkan bahwa model memiliki fit yang sempurna (*perfect fit*). Terlihat pada output yang dihasilkan berikut :

Tabel 3. Output Goodness Of Fit Metod Generalized Least Squares

Goodness Of Fit	Cut Value	Of	Nilai yang diperoleh	Kesimpulan Model
probability	$\geq 0.05$		1.000	Baik
RMSEA	$\leq 0.08$		0.000	Baik
GFI	$\geq 0.90$		0.907	Baik
AGFI	$\geq 0.90$		0.844	Tidak Baik
CMIN/DF	$\leq 2.00$		0,934	Baik
TLI	$\geq 0.95$		0.000	Tidak Baik
CFI	$\geq 0.95$		0.000	Tidak Baik

Berdasarkan hasil output diatas terlihat bahwa model yang sesuai dengan cut of value adalah probability, RMSEA, GFI, dan CMIN/DF merupakan model yang baik dan cocok dengan data.

### Analisis Signifikansi Koefisien Jalur

Analisis signifikansi koefisien jalur dapat menggunakan tabel regression weight yang ditampilkan oleh program amos. Berikut tabel regression weight yang ditampilkan:

Tabel 4. Regression Weights

			Estimate	S.E.	C.R.	P	Label
y1	<---	x	,201	,179	1,122	,262	
y2	<---	x	-,245	,180	-1,364	,173	
y2	<---	y1	,170	,180	,945	,345	
y3	<---	x	,167	,188	,891	,373	
y3	<---	y2	-,042	,185	-,229	,819	
y3	<---	y1	,259	,185	1,402	,161	

uji signifikansi koefisien jalur melalui besaran regression weight dapat dilihat dari nilai C.R. (*critical ratio*). Nilai C.R. (*critical ratio*) identik dengan nilai t-hitung dalam analisis regresi. Oleh karena itu, nilai C.R. (*critical ratio*) yang lebih besar dari 2.0 menunjukkan adanya *causal relationship* yang signifikan. Adapun jalur-jalur yang menghasilkan nilai C.R yang kurang dari sama dengan 2.0 dan nilai C.R yang lebih besar dari  $-2.0$  menandakan koefisien jalur yang dihasilkan tidak berpengaruh signifikan. Berdasarkan tabel 6 diatas, diperoleh koefisien jalur yang tidak signifikan dilihat dari nilai P yang diperoleh lebih besar dari 0.05 dan nilai C.R yang berkisar  $-0,2 < C.R < 2,0$  Hasil analisis dengan analisis jalur faktor-faktor penyebab kriminalitas di Indonesia menyatakan tidak ada jalur yang signifikan untuk mendukung teori-teori sebelumnya tentang jalur-jalur yang signifikan terhadap pengaruh langsung dan tidak langsung faktor-faktor penyebab kriminalitas di Indonesia.

### Kesimpulan

Berdasarkan penelitian yang telah dilakukan, dapat diambil beberapa kesimpulan diantaranya adalah sebagai berikut :

- 1) Jumlah kriminalitas di Indonesia dipengaruhi secara langsung oleh persentase kepadatan penduduk sebesar 0,162. Tingkat pengangguran terbuka sebesar 0,251. Persentase penduduk miskin sebesar -0,041.
- 2) Variabel-variabel yang tidak berpengaruh langsung dari faktor-faktor penyebab kriminalitas di Indonesia adalah persentase kepadatan penduduk melalui tingkat pengangguran terbuka dan persentase penduduk miskin adalah sebesar 0,059. Pengaruh tidak langsung tingkat pengangguran terbuka terhadap jumlah kriminal melalui persentase penduduk miskin adalah sebesar -0,007.

### Saran

Dari analisis, pembahasan dan kesimpulan yang diperoleh, saran-saran yang dapat disampaikan adalah sebagai berikut : Bagi peneliti selanjutnya agar dapat menambah variabel-variabel yang akan diteliti agar mendapatkan model yang lebih mewakili keadaan sebenarnya, dan bagi badan pusat statistik, mohon agar publikasi datanya lebih di spesifikkan lagi agar objek yang akan diamati dapat disajikan dengan lebih akurat.

### DAFTAR PUSTAKA

- Anonim. 2014. *Publikasi Statistika Kriminal*. Diakses Tanggal 22 Januari 2015).
- Arbuckle, J. (1997). *Amos User Guide Version 3.6*. Smallwaters Corporation: Chicago IL.
- Baumgartner, H dan Homburg, C. 1996. Applications Of Structural Equation Modeling In Marketing And Consumer Research: A review. *International Journal of Research in Marketing. Volume 13*, No.13.139-161. <http://web.cenet.org.cn/upfile/59691.pdf>. (Diakses Tanggal 12 Februari 2015).
- Dillon, W.R. & M. Golstein. 1984. *Multivariate Analysis Method & Applications*. John Wiley dan Sons, inc, Canada.
- Hair, J.F. , Anderson, R.E., Tatham, R.L., Black, W.C. 1995. *Multivariate Data Analysis*. New Jersey: Prebtice Hall.
- Hakim, 2009. *Analisis Determinan Tingkat Kejahatan Properti di Jawa Tahun 2007.*, Fakultas Ekonomi, Universitas Indonesia, Jakarta. Tidak dipublikasikan.
- Hardianto, 2009. *Analisis Faktor-Faktor yang Mempengaruhi Tingkat Kriminalitas di Indonesia dari Pendekatan Ekonomi*. Fakultas Ekonomi, Universitas Katolik Parahyangan, Bandung. Tidak dipublikasikan.

- Hulland, J., Chow, Y.H and Lam, S. 1996. "Use of Causal Models in Marketing Research : A Review". *Internatioanl Journal of Research in Marketing*. <http://webkuliah.unimedia.ac.id/ebook/files/hulland%20et%20al%281996%29%20causal%20model.pdf>. (Diakses Tanggal 22 Desember 2014).
- Kurniati, S. 2011. *Analisis Faktor-Faktor yang Mempengaruhi Tingkat Pengangguran di Indonesia (Path Analysis)*. Skripsi, Jurusan Matematika. FMIPA. Universitas Bengkulu. Bengkulu. Tidak Dipublikasikan.
- Mariana, S. dan Budiantara, I. N. 2013. Pemodelan Faktor-Faktor yang Mempengaruhi Persentase Kriminalitas di Jawa Timur dengan Pendekatan Regresi Semiparametrik Spline. *JURNAL SAINS DAN SENI POMITS Vol. 2. No. 2. (2013)*. [http://ejurnal.its.ac.id/index.php/sains\\_seni/article/viewFile/4512/1358](http://ejurnal.its.ac.id/index.php/sains_seni/article/viewFile/4512/1358). (Diakses Tanggal 17 November 2014).
- Maschiliah, S. dan suryani. E. 2012. Analisis Pengaruh Faktor-Faktor Internal dan Eksternal Penyebab Timbulnya Tindakan Criminal dengan Pendekatan Simulasi System Dinamik Untuk Mengurangi Angka Kriminalitas. *JURNAL teknik POMITS Vol.1, No.1, (2012)*. <http://digilib.its.ac.id/public/ITS-paper-27013-5208100084-Paper.pdf>. (Diakses Tanggal 17 November 2014).
- Nugroho, S. 2008. *Statistika Multivariat Terapan*, edisi I. Unib Press, Bengkulu.
- Pedhazur, E. 1982. *Multiple Regression in behavioral research*. New York : CBS Collage Publihsing.
- Putra, P. 2008. *Identifikasi Faktor-Faktor yang Mempengaruhi Tingkat Kriminal Berdasarkan Karakteristik Pelaku Kriminal dengan Menggunakan Metode Pohon Klasifikasi*. Jurusan Matematika, Universitas Andalas, Padang. Tidak dipublikasikan.
- Rahayu, S. 2013. *Aplikasi Metode Trimming pada Analisis Jalur*. [www.repository.upi.edu](http://www.repository.upi.edu). (Diakses Tanggal 27 September 2014).
- Riduwandan Engkos, K. A. 2007. *Cara Menggunakan dan Memaknai Analisis Jalur (Path Analysis)*. Alfabeta: Bandung.
- Sarwono, J. 2007. *Analisis Jalur untuk Riset Bisnis*. Andi Offset: Yogyakarta.
- Sullivan, A. 2009. *Urban Economics*, Seventh Edition. McGraw-Hill/Irwin: New York.
- Tanaka, J.S., dan Huba, G.J., 1989, A General Coefficient of Determination for Covariance Structure Models under Arbitrary GLS Estimation, *British Journal of Mathematical and Statistical Psychology* [https://books.google.co.id/books?id=jta8Y0ggiFEC&pg=PA723&lpg=PA723&dq=A+General+Coefficient+of+Determination+for+Covariance+Structure+Models+under+Arbitrary+GLS+Estimation&source=bl&ots=2Lbb2E4Vjs&sig=03M9\\_Wkpw3iY\\_8uwuHyTnSFpWc&hl=en&sa=X&ei=pR3UVIrgNYPmuQTwvoCAAg&redir\\_esc=y#v=onepage&q=A%20General%20Coefficient%20of%20Determination%20for%20Covariance%20Structure%20Models%20under%20Arbitrary%20GLS%20Estimation&f=false](https://books.google.co.id/books?id=jta8Y0ggiFEC&pg=PA723&lpg=PA723&dq=A+General+Coefficient+of+Determination+for+Covariance+Structure+Models+under+Arbitrary+GLS+Estimation&source=bl&ots=2Lbb2E4Vjs&sig=03M9_Wkpw3iY_8uwuHyTnSFpWc&hl=en&sa=X&ei=pR3UVIrgNYPmuQTwvoCAAg&redir_esc=y#v=onepage&q=A%20General%20Coefficient%20of%20Determination%20for%20Covariance%20Structure%20Models%20under%20Arbitrary%20GLS%20Estimation&f=false). (Diakses Tanggal 11 Januari 2015).
- Wibowo, A. 2005. *Pengantar Analisis Jalur (Path Analysis)*. Lembaga Penelitian Universitas Airlangga: Surabaya.
- Zhang, W. 2014. *Learning Variables In Class Laptop Multitasking and Acadmic Performance: A path analysis*. *Jurnal, Computer dan Education* 81 (2015) 82-88. <http://dx.doi.org/10.1016/j.compedu.2014.09.012>. *Compedu*. 2014. 09.012. (Diakses Tanggal 15 Januari 2015).

# ANALISIS KERAGAMAN MODEL RANCANGAN ACAK LENGKAP (RAL) DAN RANCANGAN ACAK KELOMPOK LENGKAP (RAKL) DENGAN MATRIKS RANCANGAN TERPARTISI

Fajria Mustika Ayu Wandira<sup>1</sup>, Sigit Nugroho<sup>2</sup>, Dyah Setyo Rini<sup>3</sup>

Jurusan Matematika, FMIPA, Universitas Bengkulu  
Jl. WR. Supratman Kandang Limun Bengkulu 38371A  
Email: [fajriamustikaaw@yahoo.com](mailto:fajriamustikaaw@yahoo.com)

## ABSTRAK

Penelitian ini bertujuan untuk mengkaji analisis keragaman model Rancangan Acak Lengkap (RAL) dan Rancangan Acak Kelompok Lengkap (RAKL) dengan matriks rancangan terpartisi serta mengetahui penerapannya. Matriks rancangan terpartisi merupakan solusi yang digunakan untuk melakukan perhitungan jumlah kuadrat dan derajat bebas dari masing-masing sumber keragaman.

**Kata Kunci:** Matriks Rancangan Terpartisi, Jumlah Kuadrat, Derajat Bebas.

## 1. PENDAHULUAN

Suatu percobaan merupakan suatu penelaahan ilmiah terencana yang dirancang untuk meneliti satu atau lebih populasi. Beberapa kondisi yang mencirikan suatu populasi disebut perlakuan, jadi setiap perlakuan secara khas mendefinisikan populasi [3]. Percobaan pada umumnya dilakukan untuk menemukan sesuatu. Oleh karena itu secara teoritis, percobaan diartikan sebagai tes [2] atau penyelidikan terencana untuk mendapatkan fakta baru [5]. Rancangan percobaan dapat diartikan juga sebagai tes atau serangkaian tes dimana perubahan yang berarti dilakukan pada variabel dari suatu proses atau sistem, sehingga kita dapat mengamati dan mengidentifikasi alasan-alasan perubahan pada respon output [2].

Kebanyakan rancangan percobaan dapat dikatakan sebagai modifikasi atau pengembangan dari rancangan-rancangan dasar. Rancangan dapat dikembangkan karena adanya kebutuhan untuk menangani situasi penelitian yang lebih kompleks. Rancangan percobaan yang paling sering digunakan dalam penelitian adalah Rancangan Acak Lengkap (RAL) dan Rancangan Acak Kelompok Lengkap (RAKL).

Dalam percobaan menggunakan Rancangan Acak Lengkap (RAL), sebanyak  $n$  satuan percobaan diperlukan. Satuan-satuan percobaan tersebut diusahakan seseragam mungkin kondisinya, sehingga tidak ada sumber keragaman lain yang dapat dikendalikan. Sedangkan Rancangan Acak Kelompok Lengkap (RAKL) merupakan pengembangan sederhana dari Rancangan Acak Lengkap (RAL), dengan memasukkan komponen tambahan untuk pemblokkan [3].

Salah satu tahapan dalam Analisis Keragaman rancangan percobaan adalah menghitung jumlah kuadrat dari masing-masing sumber keragaman. Dalam Rancangan Acak Lengkap (RAL), sumber-sumber keragaman tersebut adalah perlakuan atau ditulis dengan notasi  $\tau$  dan galat percobaan. Sedangkan dalam Rancangan Acak Kelompok Lengkap (RAKL), sumber-sumber keragaman tersebut adalah blok atau ditulis dengan notasi  $\beta$ , perlakuan ( $\tau$ ) dan galat percobaan. Masing-masing sumber keragaman memiliki Nilai Harapan Kuadrat Tengah (NHKT) dan derajat bebas, yang akan digunakan sebagai derajat bebas  $F$ .

Hasil percobaan umumnya diarahkan pada prosedur Analisis Keragaman atau dikenal juga dengan Analisis Varian (ANOVA). Untuk mendapatkan jumlah

kuadrat juga dapat menggunakan Dekomposisi QR. Persyaratan Dekomposisi QR menghendaki bahwa jumlah baris harus melebihi atau sama dengan jumlah kolomnya. Alternatifnya, dalam penelitian ini digunakan matriks rancangan terpartisi untuk menentukan besarnya jumlah kuadrat. Dengan menggunakan jumlah kuadrat yang telah dibahas oleh Nugroho (2014), penulis juga dapat menentukan derajat bebas, Nilai Harapan Kuadrat Tengah (NHKT), dan sebarannya.

Metode matriks rancangan terpartisi untuk perhitungan jumlah kuadrat dapat diterapkan pada jenis-jenis rancangan seperti pada Rancangan Acak Lengkap (RAL), Rancangan Acak Kelompok Lengkap (RAKL), Rancangan Persegi Latin dan Percobaan Tiga Faktor yang telah dibahas oleh Nugroho (2014), Rancangan Percobaan Tersarang Seimbang Dua dan Tiga Tahap yang telah dibahas oleh Alvionita (2015).

## 2. HASIL DAN PEMBAHASAN

### 2.1 Notasi Aljabar Matriks Rancangan Acak Lengkap (RAL)

Model linier yang digunakan pada notasi aljabar matriks pada umumnya ditulis  $Y = X\beta + \varepsilon$  dimana  $Y$  merupakan variabel acak dari vektor pengamatan berukuran  $n \times 1$ ,  $X$  merupakan matriks rancangan berukuran  $n \times p$ ,  $\beta$  merupakan vektor parameter model linier berukuran  $p \times 1$ , dan  $\varepsilon$  merupakan variabel acak dari vektor galat percobaan berukuran  $n \times 1$ . Model yang biasa digunakan pada Analisis Keragaman adalah model yang memiliki elemen matriks yang terdiri dari angka nol dan satu [1].

Model Rancangan Acak Lengkap (RAL) dalam notasi aljabar matriks dapat dituliskan menjadi :

$$Y_{rt \times 1} = X_{rt \times (t+1)} \beta_{(t+1) \times 1} + \varepsilon_{rt \times 1}$$

dengan  $Y_{rt \times 1}$  merupakan vektor pengamatan berukuran  $rt \times 1$ ,  $X_{rt \times (t+1)}$  merupakan matriks rancangan berukuran  $rt \times (t+1)$ ,  $\beta_{(t+1) \times 1}$  merupakan vektor parameter model linier berukuran  $(t+1) \times 1$ , dan  $\varepsilon_{rt \times 1}$  merupakan vektor galat percobaan berukuran  $rt \times 1$ . Matriks rancangan tersebut dipartisi menjadi beberapa sumber keragaman yang bervariasi yaitu konstanta dan perlakuan. Bentuk matriks partisi dari rancangan tersebut dapat dilihat sebagai berikut:

$$Y_{rt \times 1} = \begin{bmatrix} Y_{11} \\ \vdots \\ Y_{1r} \\ Y_{21} \\ \vdots \\ Y_{2r} \\ \vdots \\ Y_{t1} \\ \vdots \\ Y_{t1} \\ \vdots \\ Y_{tr} \\ \vdots \\ Y_{tr} \\ \vdots \\ Y_{t1} \\ \vdots \\ Y_{tr} \\ \vdots \\ Y_{tr} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \vdots \\ \tau_t \end{bmatrix}$$

$$Y_{rt \times 1} \quad [1_{r \times 1} \otimes 1_{t \times 1} + I_{t \times t} \otimes 1_{r \times 1}]$$

Matriks partisi terdiri dari sumber keragaman konstanta  $X_\mu = [1_{r \times 1} \otimes 1_{t \times 1}]$  dan sumber keragaman perlakuan  $X_\tau = [I_{t \times t} \otimes 1_{r \times 1}]$ .

Matriks proyeksi memiliki bentuk  $M = X(X^t X)^{-1} X^t$ . Selanjutnya, matriks rancangan yang sudah di partisi diuji dengan menggunakan matriks proyeksi. Matriks proyeksi yang dihasilkan adalah sebagai berikut:

$$M_\mu = \frac{1}{rt} J_{r \times r} \otimes J_{t \times t}$$

$$M_{\tau} = \frac{1}{r} I_{t \times t} \otimes J_{r \times r}$$

Selanjutnya, dengan menggunakan sifat *Cronecker Product* (Perkalian Kronecker), penulis dapat menemukan setiap kombinasi perkalian matriks dari  $M_{\mu}$  dan  $M_{\tau}$ . Berikut adalah hasil perkalian matriks dari  $M_{\mu}$  dan  $M_{\tau}$ :

Tabel 1. Perkalian matriks

	$M_{\mu}$	$M_{\tau}$
$M_{\mu}$	$M_{\mu}$	$M_{\mu}$
$M_{\tau}$	$M_{\mu}$	$M_{\tau}$

### 2.1.1 Notasi Matriks pada Jumlah Kuadrat dan Derajat Bebas

Pada umumnya formula jumlah kuadrat dapat ditulis dalam bentuk notasi matriks sebagai berikut (lihat Nugroho (2014)):

- JK Total =  $Y^t(I - M_{\mu})Y$
- JK Perlakuan =  $Y^t(M_{\tau} - M_{\mu})Y$
- JK Error =  $Y^t(I - M_{\tau})Y$

Bahwa  $Y$  vektor acak berukuran  $n \times 1$  berdistribusi  $N(y; \mu, I)$ . Variabel acak  $U = Y^t A Y$  berdistribusi  $\chi^2(u; K; \lambda)$ , dimana  $\lambda = \mu^t A \mu / 2$ , jika dan hanya jika  $A$  matriks idempotent dengan  $\text{rank}(A) = K$  [1]. Dari sifat simetri dan idempoten, maka rank nya adalah:

$$\begin{aligned} \text{rank}(I - M_{\mu}) &= rt - 1 \\ \text{rank}(M_{\tau} - M_{\mu}) &= t - 1 \\ \text{rank}(I - M_{\tau}) &= (r - 1)t \end{aligned}$$

### 2.1.2 Statistik Uji Rancangan Acak Lengkap (RAL)

Matriks  $B$  dikalikan matriks  $A$  sama dengan nol apabila matriks  $A$  dan matriks  $B$  saling bebas. Akan dibuktikan bentuk kuadrat  $Y^t(M_{\tau} - M_{\mu})Y$  dan  $Y^t(I - M_{\tau})Y$  saling bebas:

$$\begin{aligned} (M_{\tau} - M_{\mu})(I - M_{\tau}) &= M_{\tau} - M_{\tau} \cdot M_{\tau} - M_{\mu} + M_{\mu} \cdot M_{\tau} \\ &= M_{\tau} - M_{\tau} - M_{\mu} + M_{\mu} \\ &= 0 \end{aligned}$$

karena  $(M_{\tau} - M_{\mu})(I - M_{\tau}) = 0$ , maka kedua bentuk kuadrat  $Y^t(M_{\tau} - M_{\mu})Y$  dan  $Y^t(I - M_{\tau})Y$  saling bebas. Sehingga, menurut Graybill (1976)

$$F_{hit} = \frac{KT[P]}{KT[G]} = \frac{\frac{Y^t(M_{\tau} - M_{\mu})Y}{t - 1}}{\frac{Y^t(I - M_{\tau})Y}{(r - 1)t}}$$

menyebar menurut sebaran  $F$  dengan derajat bebas  $t - 1$  dan  $(r - 1)t$ .

## 2.2 Notasi Aljabar Matriks Rancangan Acak Kelompok Lengkap (RAKL)

$$Y_{bt \times 1} = X_{bt \times (b+t+1)} \beta_{(b+t+1) \times 1} + \epsilon_{bt \times 1}$$

dengan  $Y_{bt \times 1}$  merupakan vektor pengamatan berukuran  $bt \times 1$ ,  $X_{bt \times (b+t+1)}$  merupakan matriks rancangan berukuran  $bt \times (b + t + 1)$ ,  $\beta_{(b+t+1) \times 1}$  merupakan vektor parameter model linier berukuran  $(b + t + 1) \times 1$ , dan  $\epsilon_{bt \times 1}$  merupakan vektor galat percobaan berukuran  $bt \times 1$ . Matriks rancangan tersebut dipartisi menjadi beberapa sumber keragaman yang bervariasi yaitu konstanta, blok, dan perlakuan. Bentuk matriks partisi dari rancangan tersebut dapat dilihat sebagai berikut:

$$\begin{matrix} Y_{11} \\ \vdots \\ Y_{1t} \\ Y_{21} \\ \vdots \\ Y_{2t} \\ \vdots \\ Y_{b1} \\ \vdots \\ Y_{.t} \end{matrix} = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 1 & \cdots & 1 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{matrix} \mu \\ \beta_1 \\ \vdots \\ \beta_b \\ \tau_1 \\ \vdots \\ \tau_t \end{matrix}$$

$$\mathbf{Y}_{rt \times 1} = [\mathbf{1}_{b \times 1} \otimes \mathbf{1}_{t \times 1} \mid \mathbf{I}_{b \times b} \otimes \mathbf{1}_{t \times 1} \mid \mathbf{1}_{b \times 1} \otimes \mathbf{I}_{t \times t}]$$

Matriks partisi terdiri dari sumber keragaman konstanta  $\mathbf{X}_\mu = [\mathbf{1}_{b \times 1} \otimes \mathbf{1}_{t \times 1}]$ , sumber keragaman blok  $\mathbf{M}_\beta = [\mathbf{I}_{b \times b} \otimes \mathbf{1}_{t \times 1}]$ , dan sumber keragaman perlakuan  $\mathbf{X}_\tau = [\mathbf{1}_{b \times 1} \otimes \mathbf{I}_{t \times t}]$

Matriks proyeksi memiliki bentuk  $\mathbf{M} = \mathbf{X}(\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t$ . Selanjutnya, matriks rancangan yang sudah di partisi diuji dengan menggunakan matriks proyeksi. Matriks proyeksi yang dihasilkan adalah sebagai berikut:

$$\begin{aligned} \mathbf{M}_\mu &= \frac{1}{bt} \mathbf{J}_{b \times b} \otimes \mathbf{J}_{t \times t} \\ \mathbf{M}_\beta &= \frac{1}{t} \mathbf{I}_{b \times b} \otimes \mathbf{J}_{t \times t} \\ \mathbf{M}_\tau &= \frac{1}{b} \mathbf{J}_{b \times b} \otimes \mathbf{I}_{t \times t} \end{aligned}$$

Selanjutnya, dengan menggunakan sifat *Cronecker Product* (Perkalian Kronecker), penulis dapat menemukan setiap kombinasi perkalian matriks dari  $\mathbf{M}_\mu$ ,  $\mathbf{M}_\beta$  dan  $\mathbf{M}_\tau$ . Berikut adalah hasil perkalian matriks dari  $\mathbf{M}_\mu$ ,  $\mathbf{M}_\beta$  dan  $\mathbf{M}_\tau$ :

Tabel 2. Perkalian matriks

	$\mathbf{M}_\mu$	$\mathbf{M}_\beta$	$\mathbf{M}_\tau$
$\mathbf{M}_\mu$	$\mathbf{M}_\mu$	$\mathbf{M}_\mu$	$\mathbf{M}_\mu$
$\mathbf{M}_\beta$	$\mathbf{M}_\mu$	$\mathbf{M}_\beta$	$\mathbf{M}_\mu$
$\mathbf{M}_\tau$	$\mathbf{M}_\mu$	$\mathbf{M}_\mu$	$\mathbf{M}_\tau$

### 2.2.1 Notasi Matriks pada Jumlah Kuadrat dan Derajat Bebas

Formula jumlah kuadrat dapat ditulis dalam bentuk notasi matriks sebagai berikut (lihat Nugroho (2014)):

- JK Total =  $\mathbf{Y}^t (\mathbf{I} - \mathbf{M}_\mu) \mathbf{Y}$
- JK Blok =  $\mathbf{Y}^t (\mathbf{M}_\beta - \mathbf{M}_\mu) \mathbf{Y}$
- JK Perlakuan =  $\mathbf{Y}^t (\mathbf{M}_\tau - \mathbf{M}_\mu) \mathbf{Y}$
- JK Error =  $\mathbf{Y}^t (\mathbf{I} + \mathbf{M}_\mu - \mathbf{M}_\tau - \mathbf{M}_\beta) \mathbf{Y}$

$\mathbf{Y}$  vektor acak berukuran  $n \times 1$  berdistribusi  $\mathbf{N}(\mathbf{y}; \boldsymbol{\mu}, \mathbf{I})$ . Variabel acak  $\mathbf{U} = \mathbf{Y}^t \mathbf{A} \mathbf{Y}$  berdistribusi  $\chi^2(u; K; \lambda)$ , dimana  $\lambda = \boldsymbol{\mu}^t \mathbf{A} \boldsymbol{\mu} / 2$ , jika dan hanya jika  $\mathbf{A}$  matriks idempotent dengan  $\text{rank}(\mathbf{A}) = K$  [1]. Dari sifat simetri dan idempoten, maka rank nya adalah:

$$\begin{aligned} \text{rank}(\mathbf{I} - \mathbf{M}_\mu) &= bt - 1 \\ \text{rank}(\mathbf{M}_\beta - \mathbf{M}_\mu) &= b - 1 \end{aligned}$$

$$\begin{aligned} \text{rank}(M_\tau - M_\mu) &= t - 1 \\ \text{rank}(I + M_\mu - M_\tau - M_\beta) &= (b - 1)(t - 1) \end{aligned}$$

### 2.2.2 Statistik Uji Rancangan Acak Kelompok Lengkap (RAKL)

Peneliti akan menunjukkan bahwa bentuk kuadratsaling bebas:

- a. Bentuk kuadrat  $Y^t(M_\tau - M_\mu)Y$  dan  $Y^t(I + M_\mu - M_\tau - M_\beta)Y$  saling bebas  
Akan dibuktikan bentuk kuadrat  $Y^t(M_\tau - M_\mu)Y$  dan  $Y^t(I + M_\mu - M_\tau - M_\beta)Y$  saling bebas:

$$\begin{aligned} &(M_\tau - M_\mu)(I + M_\mu - M_\tau - M_\beta) \\ &= M_\tau + M_\tau M_\mu - M_\tau M_\tau - M_\tau M_\beta - M_\mu - M_\mu M_\mu + M_\mu M_\tau + M_\mu M_\beta \\ &= M_\tau + M_\mu - M_\tau - M_\mu - M_\mu - M_\mu + M_\mu + M_\mu \\ &= 0 \end{aligned}$$

karena  $(M_\tau - M_\mu)(I + M_\mu - M_\tau - M_\beta) = 0$ , maka berdasarkan Teorema 2.9 kedua bentuk kuadrat  $Y^t(M_\tau - M_\mu)Y$  dan  $Y^t(I + M_\mu - M_\tau - M_\beta)Y$  saling bebas. Sehingga, menurut Graybill (1976)

$$F_{hit} = \frac{KT[P]}{KT[G]} = \frac{\frac{Y^t(M_\tau - M_\mu)Y}{t - 1}}{\frac{Y^t(I + M_\mu - M_\tau - M_\beta)Y}{(b - 1)(t - 1)}}$$

menyebar menurut sebaran  $F$  dengan derajat bebas  $t - 1$  dan  $(b - 1)(t - 1)$ .

- b. Bentuk kuadrat  $Y^t(M_\beta - M_\mu)Y$  dan  $Y^t(I + M_\mu - M_\tau - M_\beta)Y$  saling bebas  
Akan dibuktikan bentuk kuadrat  $Y^t(M_\beta - M_\mu)Y$  dan  $Y^t(I + M_\mu - M_\tau - M_\beta)Y$  saling bebas:

$$\begin{aligned} &(M_\beta - M_\mu)(I + M_\mu - M_\tau - M_\beta) \\ &= M_\beta + M_\beta M_\mu - M_\beta M_\tau - M_\beta M_\beta - M_\mu - M_\mu M_\mu + M_\mu M_\tau + M_\mu M_\beta \\ &= M_\beta + M_\mu - M_\mu - M_\beta - M_\mu - M_\mu + M_\mu + M_\mu \\ &= 0 \end{aligned}$$

karena  $(M_\beta - M_\mu)(I + M_\mu - M_\tau - M_\beta) = 0$ , maka berdasarkan Teorema 2.9 kedua bentuk kuadrat  $Y^t(M_\beta - M_\mu)Y$  dan  $Y^t(I + M_\mu - M_\tau - M_\beta)Y$  saling bebas. Sehingga, menurut Graybill (1976)

$$F_{hit} = \frac{KT[P]}{KT[G]} = \frac{\frac{Y^t(M_\beta - M_\mu)Y}{t - 1}}{\frac{Y^t(I + M_\mu - M_\tau - M_\beta)Y}{(b - 1)(t - 1)}}$$

menyebar menurut sebaran  $F$  dengan derajat bebas  $t - 1$  dan  $(b - 1)$

### 3. TELADAN PENERAPAN

Teladan yang digunakan pada Rancangan Acak Lengkap (RAL) dan Rancangan Acak Kelompok Lengkap (RAKL) diambil dari teladan pada buku "Design and Analysis of Experiment" oleh Angela Dean dan Daniel Voss (1999).

#### 3.1 Teladan Penerapan Rancangan Acak Lengkap (RAL)

Tabel 3. Data Untuk Percobaan Ikan Air Tawar

Kode	Hemoglobin (gram per 100 ml)									
1	6.7	7.8	5.5	8.4	7.0	7.8	8.6	7.4	5.8	7.0
2	9.9	8.4	10.4	9.3	10.7	11.9	7.1	6.4	8.6	10.6
3	10.4	8.1	10.6	8.7	10.7	9.1	8.8	8.1	7.8	8.0
4	9.3	9.3	7.2	7.8	9.3	10.2	8.7	8.6	9.3	7.2

Nilai perhitungan yang diperoleh dengan menggunakan Program R, dapat disajikan pada tabel di bawah ini:

Tabel 4. Analisis Keragaman untuk Teladan Penerapan Rancangan Acak Lengkap (RAL)

Sumber Keragaman	<i>db</i>	<i>JK</i>	<i>KT</i>	$F_{hit}$	$F_{tabel}$
Perlakuan	3	26,803	8,934	5,694	2.866
Galat	36	56,471	1.569		
Total	39	83.274			

Berdasarkan Tabel 4, diperoleh  $F_{hit} = 5,694 > F_{tabel} = 2.866$ , sehingga  $H_0$  ditolak. Artinya terdapat perbedaan yang signifikan antara ikan air tawar yang diberikan makanan dengan kadar yang berbeda terhadap hasil hemoglobin.

#### 3.1 Teladan Penerapan Rancangan Acak Kelompok Lengkap (RAKL)

Tabel 5. Data Penatalaksanaan (Protokol) Rawat Inap yang Diabaikan

Subjek	Protokol		
	1	2	3
1	7131	6846	7095
2	8062	8573	8685
3	6921	7287	7132
4	7249	7554	7471
5	9551	8866	8840
6	7046	7681	6939
7	7715	7535	7831
8	9862	10087	9711
9	7812	7708	8179

Nilai perhitungan yang diperoleh dengan menggunakan Program R, dapat disajikan pada tabel di bawah ini:

Tabel 6. Analisis Keragaman untuk Teladan Penerapan Rancangan Acak Kelompok Lengkap (RAKL)

Sumber Keragaman	db	JK	KT	$F_{hit}$	$F_{tabel}$
Blok	8	23117462,296	2889682,787	37,423	3,634
Perlakuan	2	35948,741	17974,371	0,233	
Galat	16	1235483,259	77217,704		
Total	26	24388894,296			

Berdasarkan Tabel 6, pengujian pertama dilakukan untuk mengetahui pengaruh blok terhadap respon, diperoleh  $F_{hit} = 37,423 > F_{tabel} = 3,634$ , sehingga  $H_0$  ditolak. Artinya terdapat perbedaan yang signifikan antara subjek terhadap efek penatalaksanaan (protokol) yang diabaikan. Untuk pengujian kedua dilakukan untuk mengetahui pengaruh protokol terhadap efek penatalaksanaan (protokol) yang diabaikan, diperoleh  $F_{hit} = 0,233 < F_{tabel} = 3,634$ , sehingga  $H_0$  ditolak. Artinya tidak terdapat perbedaan yang signifikan antara protokol terhadap efek penatalaksanaan (protokol) yang diabaikan.

#### 4. DAFTAR PUSTAKA

- [1] Graybill, F. A. 1976. *Theory and Application of the Linier Model*. Wadsworth
- [2] Montgomery, D. C. 2001. *Design and Analysis of Experiments*. Fifth Edition. John Wiley & Sons. New York.
- [3] Nugroho, S. 2008. *Dasar-dasar Rancangan Percobaan*. UNIB Press. Bengkulu.
- [4] Nugroho, S. 2014. Metode Matriks Rancangan Terpartisi untuk Perhitungan Jumlah Kuadrat. *Prosiding Semirata BKS PTN Wilayah Barat*. Bogor.
- [5] Steel, R.G.D dan J.H, Torrie. 1995. *Prinsip dan Prosedur Statistika*. Penerjemah Bambang Sumantri. Gramedia Pustaka, Jakarta.

# ANALISIS KLASTER HIERARKI DIVISIF PADA KABUPATEN/KOTA DI PROVINSI BENGKULU BERDASARKAN TINGKAT PENDIDIKAN TERTINGGI TAHUN 2015

## (ANALYSIS OF DIVISIVE HIERARCHICAL CLUSTERING OF BENGKULU DISTRICTS BASED ON HIGHEST EDUCATION LEVEL IN 2015)

Riski Alawiyah<sup>1\*</sup>, Sigit Nugroho<sup>2</sup>, Fachri Faisal<sup>3</sup>  
Jurusan Matematika, FMIPA, Universitas Bengkulu  
Jl. WR. Supratman Kandang Limun Bengkulu 38371A  
e-mail : [riskialawiyah53@gmail.com](mailto:riskialawiyah53@gmail.com)<sup>1\*</sup>

### ABSTRACT

This study aims to obtain a deviant hierarchy cluster solution from the highest education data of the population aged above 15 years for Bengkulu Districts in 2015. Analysis of divisive hierarchical clustering is one of the hierarchical methods that perform clustering processes in the opposite direction of the agglomerative method. The process starts from one main cluster, then separates the objects from the main cluster to form a splinter group. The process of divisive hierarchical clustering analysis involves an inequality matrix containing Euclidean distances between two objects. Based on data, there is 6 splinter groups. The splinter groups are depicted in a dendrogram. Based on the similarity size on the dendrogram, the author take 2 or 3 cluster separations.

**Keywords :** *euclidean distance, divisive hierarchical cluster, splinter groups.*

### ABSTRAK

Penelitian ini bertujuan untuk memperoleh solusi klaster hierarki divisif dari data pendidikan tertinggi penduduk yang berumur di atas 15 tahun untuk setiap kabupaten/kota di Provinsi Bengkulu pada tahun 2015. Analisis klaster hierarki divisif adalah salah satu metode hierarki yang melakukan proses pengklasteran dengan arah yang berlawanan dari metode aglomeratif. Proses dimulai dari satu klaster utama, kemudian melakukan pemisahan obyek-obyek dari klaster utama sehingga membentuk *splinter group*. Proses analisis klaster hierarki divisif akan melibatkan matriks ketidaksamaan yang memuat jarak *Euclid* diantara dua obyek. Berdasarkan data tersebut, diperoleh 6 *splinter group*. *Splinter group* tersebut digambarkan ke dalam sebuah dendrogram. Berdasarkan ukuran kemiripan pada dendrogram tersebut penulis mengambil 2 atau 3 pemisahan klaster.

**Kata kunci:** *jarak euclid, klaster hierarki divisif, splinter group.*

## A. PENDAHULUAN

### 1. Latar Belakang

Berdasarkan data Survei Sosial Ekonomi Nasional (SUSENAS) 2015 yaitu persentase penduduk berumur 15 tahun ke atas dari setiap kabupaten/kota di Provinsi Bengkulu untuk penduduk yang tidak sekolah mencapai 17,48%, penduduk dengan tingkat pendidikan Sekolah Dasar (SD) mencapai 25,38%, Sekolah Menengah Pertama (SMP) 22,36%, Sekolah Menengah Atas (SMA) 25,03%, Diploma 2,26% dan Sarjana 7,49%. Berdasarkan hasil survei tersebut angka persentase penduduk yang tidak sekolah cukup besar dibandingkan dengan angka persentase untuk tamatan Diploma dan Sarjana yang masih tergolong rendah [1].

Menurut [2], secara umum ada beberapa faktor penyebab rendahnya tingkat pendidikan tertinggi suatu masyarakat diantaranya adalah pernikahan dini, rendahnya ekonomi keluarga serta cara pandang orang tua mengenai pendidikan tinggi tidaklah penting. Oleh karena faktor-faktor penyebab tersebut tingkat pendidikan tertinggi masyarakat di setiap kabupaten/kota menjadi berbeda-beda. Untuk mengetahui perbedaan tingkat pendidikan tertinggi dari setiap kabupaten/kota dapat dilakukan suatu analisis kluster. Menurut [5], analisis kluster adalah suatu koleksi metode statistik yang mengidentifikasi kelompok sampel berdasarkan karakteristik yang serupa.

Menurut [6], ciri-ciri dari kluster yang baik yaitu mempunyai kesamaan (homogenitas) yang tinggi antar obyek dalam satu kluster (*within cluster*) dan mempunyai ketaksamaan (heterogenitas) yang tinggi antar kluster (*between cluster*). Jadi setiap obyek yang memiliki kesamaan paling dekat dengan obyek yang lain akan berada pada kluster yang sama.

Menurut [3], konsep dasar dari analisis kluster adalah ukuran kedekatan antar obyek yang dapat dilakukan dengan cara menghitung jarak diantara dua obyek. Tujuannya adalah untuk memperoleh matriks ketaksamaan yang memuat semua jarak diantara dua obyek. Menurut [4], pengukuran jarak yang paling umum dalam analisis kluster hierarki devisif adalah pengukuran jarak *Euclid*. Rumus persamaan untuk menghitung jarak *Euclid* adalah

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_p - y_p)^2} \quad (1)$$

dengan keterangan  $d(x, y)$  adalah jarak diantara obyek  $x$  dan  $y$ ;  $x_i$  dan  $y_i$  adalah nilai-nilai untuk obyek  $x$  dan  $y$  pada variabel ke- $i$  dan  $p$  adalah jumlah variabel.

Metode hierarki terdiri dari dua jenis yaitu aglomeratif dan devisif. Kedua metode tersebut membentuk hierarki dalam arah yang berlawanan. Berbagai penelitian telah banyak menggunakan metode aglomeratif. Sedangkan metode devisif masih sedikit. Oleh

karena itu penulis berinisiatif untuk membahas metode devisif ini, karena setelah membaca literatur metode devisif ini ternyata sangat menarik untuk dijadikan topik penelitian.

Menurut [4], metode pengklasteran hierarki devisif dimulai dengan adanya satu klaster yang memuat semua  $n$  obyek. Kemudian membagi  $n$  obyek menjadi dua kelompok. Selanjutnya tahapan pengklasteran dilakukan dengan cara yang tidak sama dengan metode aglomeratif yaitu melakukan pemisahan obyek-obyek dari klaster utama sehingga membentuk  $n$  klaster yang disebut sebagai *splinter group*. Tahapan pemisahan obyek dari klaster utama dalam penelitian devisif ini dilakukan sebanyak  $(n - 1)$  langkah.

Proses dalam analisis klaster hierarki devisif didasarkan pada perhitungan rata-rata obyek terhadap obyek lainnya. Perhitungan rata-rata tersebut dilakukan setelah diperoleh matriks ketaksamaan jarak *Euclid*. Rumus persamaan yang digunakan adalah

$$\bar{x}_k = \frac{1}{n-1} \sum_{k=1}^n x_k \quad (2)$$

dengan  $\bar{x}_k$  adalah rata-rata obyek ke- $k$ ,  $x_k$  adalah nilai-nilai pada obyek  $k$  ( $k = 1, 2, \dots, n$ ) dan  $n - 1$  adalah jumlah semua obyek dikurangi satu, karena dalam matriks ketaksamaan memuat jarak diantara obyek yang sama dengan nilai nol.

Berdasarkan penjelasan tentang metode hierarki devisif di atas, penulis tertarik untuk melakukan penelitian tentang analisis klaster hierarki devisif pada data tingkat pendidikan tertinggi masyarakat untuk setiap kabupaten/kota di Provinsi Bengkulu tahun 2015. Serta alternatif pengklasteran akan menggunakan program Microsoft Office Excel 2007.

## **2. Rumusan Masalah**

Rumusan masalah yang akan dibahas pada penelitian ini adalah bagaimana hasil analisis klaster hierarki devisif untuk semua kabupaten/kota di Provinsi Bengkulu berdasarkan tingkat pendidikan tertinggi pada tahun 2015?

## **3. Tujuan Penelitian**

Tujuan penelitian ini adalah untuk memperoleh solusi klaster hierarki devisif untuk semua kabupaten/kota di Provinsi Bengkulu berdasarkan tingkat pendidikan tertinggi pada tahun 2015.

## B. METODE PENELITIAN

Data yang digunakan dalam penelitian ini adalah persentase penduduk berumur 15 tahun ke atas menurut kabupaten/kota berdasarkan tingkat pendidikan tertinggi di Provinsi Bengkulu tahun 2015.

**Tabel 1. Obyek dan Variabel bebas dalam penelitian**

Obyek		Variabel pendidikan tertinggi	
1. Bengkulu Selatan	(BS)	1. Variabel 'tidak sekolah' berasal dari data persentase yang tidak sekolah/tidak tamat SD.	(X <sub>1</sub> )
2. Rejang Lebong	(RL)	2. Variabel 'pendidikan dasar' berasal dari data persentase tamatan SD dan SMP.	(X <sub>2</sub> )
3. Bengkulu Utara	(BU)	3. Variabel 'pendidikan menengah' berasal dari data persentase tamatan SMA	(X <sub>3</sub> )
4. Kaur	(Ka)	4. Variabel 'pendidikan tinggi' berasal dari data persentase tamatan Diploma dan Sarjana	(X <sub>4</sub> )
5. Seluma	(Se)		
6. Mukomuko	(Mm)		
7. Lebong	(Le)		
8. Kepahiang	(Ke)		
9. Bengkulu Tengah	(BT)		
10. Kota Bengkulu	(KB)		

Tahapan analisis kluster dalam penelitian ini adalah sebagai berikut:

1. Menentukan variabel penelitian dan input data ke dalam tabel pengamatan.
2. Melakukan uji asumsi dalam analisis kluster seperti deskripsi statistik sampel dan uji multikolinearitas.
3. Menghitung matriks ketaksamaan dari pengukuran jarak *Euclid*.
4. Melakukan analisis kluster hierarki devisif dengan Microsoft Office Excel 2007.
5. Membuat dendrogram dari hasil pengklasteran hierarki divisif.
6. Menentukan jumlah kluster dari hasil pengklasteran hierarki divisif.

## C. HASIL DAN PEMBAHASAN

### 1. Deskripsi Data

Tabel data yang akan digunakan dalam analisis kluster devisif ini adalah

**Tabel 2. Data tingkat pendidikan tertinggi tahun 2015**

Kabupaten/Kota	Tidak Sekolah(%)	Pendidikan Dasar(%)	Pendidikan Menengah(%)	Pendidikan Tinggi(%)
Bengkulu Selatan	14,8	47,39	27,5	10,32
Rejang Lebong	18,53	49,1	25,13	7,25
Bengkulu Utara	19,84	52,47	20,82	6,87
Kaur	21,08	47,22	22,74	8,96
Seluma	20,51	55,28	19,37	4,84
Mukomuko	21,27	54,36	16,31	8,07
Lebong	20,08	52,74	20,31	6,87
Kepahiang	22,47	49,09	22,38	6,06
Bengkulu Tengah	27,14	50,47	17,76	4,63
Kota Bengkulu	5,92	33,48	39,63	20,97

Hasil deskripsi statistik dari program IBM Statistic SPSS 21 dapat dilihat pada Tabel 3 berikut:

**Tabel 3. Deskripsi statistik dari variabel pendidikan tertinggi**

Variabel	N	Min	Maks	Rata-rata	Simpangan Baku	Varians
Tidak Sekolah	10	5,92	27,14	19,16	5,58	31,18
Pendidikan Dasar	10	33,48	55,28	49,16	6,17	38,08
Pendidikan Menengah	10	16,31	39,63	23,20	6,66	44,31
Pendidikan Tinggi	10	4,63	20,97	8,48	4,72	22,27
Nilai valid	10					

Diketahui dari Tabel 3 nilai rata-rata yang paling rendah menunjukkan persentase penduduk dengan pendidikan diploma dan sarjana di Provinsi Bengkulu sebesar 8,48%. Sedangkan nilai rata-rata yang paling tinggi menunjukkan persentase penduduk di Provinsi Bengkulu yang pendidikan tertingginya sebatas pendidikan dasar SD dan SMP sebesar 49,16%.

## 2. Pengujian Asumsi Data

Penelitian dalam analisis kluster harus fokus pada dua asumsi yaitu keterwakilan sampel dan multikolinearitas. Asumsi keterwakilan sampel telah terpenuhi karena riset penelitian ini merupakan riset populasi yang menggunakan data SUSENAS tahun 2016 mengenai persentase penduduk dari setiap kabupaten/kota di Provinsi Bengkulu berdasarkan tingkat pendidikan tertinggi tahun 2015.

Pengujian multikolinearitas yang telah dilakukan pada program IBM Statistic SPSS 21 memberikan nilai korelasi yang kuat. Namun penelitian akan tetap dilanjutkan, dengan menggunakan data yang ada, karena melihat nilai korelasi masih di bawah 1, serta tujuan dari penulis adalah untuk mendeskripsikan dan menggambarkan data tersebut melalui analisis kluster hierarki devisif.

## 3. Perhitungan Matriks Ketaksamaan Jarak *Euclid*

Pengukuran jarak digunakan dalam analisis kluster hierarki devisif adalah jarak *Euclid* yang menggunakan Persamaan (1). Contoh perhitungan manual jarak *Euclid* diantara obyek *BS* dengan obyek *RL* adalah sebagai berikut:

$$\begin{aligned}
 d(BS, RL) &= \sqrt{(BS_1 - RL_1)^2 + (BS_2 - RL_2)^2 + (BS_3 - RL_3)^2 + (BS_4 - RL_4)^2} \\
 &= \sqrt{(14,80 - 18,53)^2 + (47,39 - 49,10)^2 + (27,50 - 25,13)^2 + (10,32 - 7,25)^2} \\
 &= \sqrt{(-3,73)^2 + (-1,71)^2 + (2,37)^2 + (3,07)^2} \\
 &= \sqrt{13,91 + 2,92 + 5,62 + 9,42} \\
 &= \sqrt{31,88} \\
 &= 5,65
 \end{aligned}$$

Hasil perhitungan dari program Microsoft Office Excel 2007 sebagai berikut.

**Tabel 4. Matriks ketaksamaan jarak *Euclid***

Obyek	Jarak <i>Euclid</i>									
	<i>BS</i>	<i>RL</i>	<i>BU</i>	<i>Ka</i>	<i>Se</i>	<i>Mm</i>	<i>Le</i>	<i>Ke</i>	<i>BT</i>	<i>KB</i>
<i>BS</i>	0,00	5,65	10,38	8,00	13,82	14,86	10,96	10,30	17,00	23,08
<i>RL</i>	5,65	0,00	5,64	4,32	9,01	10,66	6,25	4,95	11,71	28,31
<i>BU</i>	10,38	5,64	0,00	6,10	3,82	5,23	0,62	4,63	8,47	33,27
<i>Ka</i>	8,00	4,32	6,10	0,00	9,68	9,65	6,46	3,74	9,53	29,12
<i>Se</i>	13,82	9,01	3,82	9,68	0,00	4,61	3,41	7,26	8,35	36,86
<i>Mm</i>	14,86	10,66	5,23	9,65	4,61	0,00	4,63	8,37	7,97	37,17
<i>Le</i>	10,96	6,25	0,62	6,46	3,41	4,63	0,00	4,90	8,16	33,82
<i>Ke</i>	10,30	4,95	4,63	3,74	7,26	8,37	4,90	0,00	6,86	32,21
<i>BT</i>	17,00	11,71	8,47	9,53	8,35	7,97	8,16	6,86	0,00	38,53
<i>KB</i>	23,08	28,31	33,27	29,12	36,86	37,17	33,82	32,21	38,53	0,00

Matriks ketaksamaan jarak *Euclid* di atas akan digunakan untuk memulai proses pengklasteran devisif. Pada setiap langkah pengklasteran, matriks ketaksamaan tersebut harus direvisi dengan mengeluarkan obyek yang telah terbentuk *splinter group*.

#### 4. Perhitungan Jarak Rata-Rata dari Matriks Ketaksamaan Jarak *Euclid*

Perhitungan rata-rata dari matriks ketaksamaan pada Tabel 4 dapat menggunakan rumus Persamaan (2). Contoh perhitungan rata-rata obyek *BS* terhadap obyek lainnya adalah sebagai berikut:

$$\begin{aligned}
 \bar{d}_{BS} &= \frac{1}{9} \sum_{i=1}^9 d_{i,1} \\
 &= \frac{1}{9} (d_{2,1} + d_{3,1} + d_{4,1} + d_{5,1} + d_{6,1} + d_{7,1} + d_{8,1} + d_{9,1} + d_{10,1}) \\
 &= \frac{1}{9} (5,65 + 10,38 + 8 + 13,82 + 14,86 + 10,96 + 10,30 + 17 + 23,09) \\
 &= \frac{1}{9} (114,06) \\
 &= 12,67
 \end{aligned}$$

Hasil perhitungan jarak rata-rata dari setiap obyek dapat dilihat pada Tabel 5.

**Tabel 5. Jarak rata-rata untuk setiap obyek terhadap obyek lainnya**

Obyek	Jarak rata-rata terhadap obyek lain
<i>BS</i>	12,67
<i>RL</i>	9,61
<i>BU</i>	8,68
<i>Ka</i>	9,62
<i>Se</i>	10,76
<i>Mm</i>	11,46
<i>Le</i>	8,80
<i>Ke</i>	9,25
<i>BT</i>	12,95
<i>KB</i>	32,49

## 5. Proses dan Hasil Analisis Kluster Hierarki Devisif

Hasil rekapitulasi dari proses pengklasteran hierarki divisif pada Microsoft Office Excel 2007 dapat dilihat pada Tabel 6.

**Tabel 6. Hasil analisis kluster hierarki devisif dengan jarak *Euclid***

No	Pemisahan obyek		Jarak kedekatan	Jarak pemisahan
	Kluster utama	SG		
0	<i>BS, RL, BU, Ka, Se, Mm, Le, Ke, BT, KB</i>	0	-	-
1	<i>BS, RL, BU, Ka, Se, Mm, Le, Ke, BT</i>	<i>KB</i>	-	32,49
2	<i>RL, BU, Ka, Se, Mm, Le, Ke, BT</i>	<i>BS</i>	-	11,37
3	<i>BU, Ka, Se, Mm, Le, Ke, BT</i>	<i>RL</i>	1,86	9,51
4	<i>BU, Se, Mm, Le, Ke, BT</i>	<i>Ka</i>	1,37	9,07
5	<i>BU, Se, Mm, Le, BT</i>	<i>Ke</i>	0,08	8,99
6	<i>BU, Se, Mm, Le</i>	<i>BT</i>	-	8,24
7	<i>BU, Se, Le</i>	<i>Mm</i>	-	4,83
8	<i>BU, Le</i>	<i>Se</i>	-	3,61
9	0	<i>BU</i>	-	0,63
		<i>Le</i>		

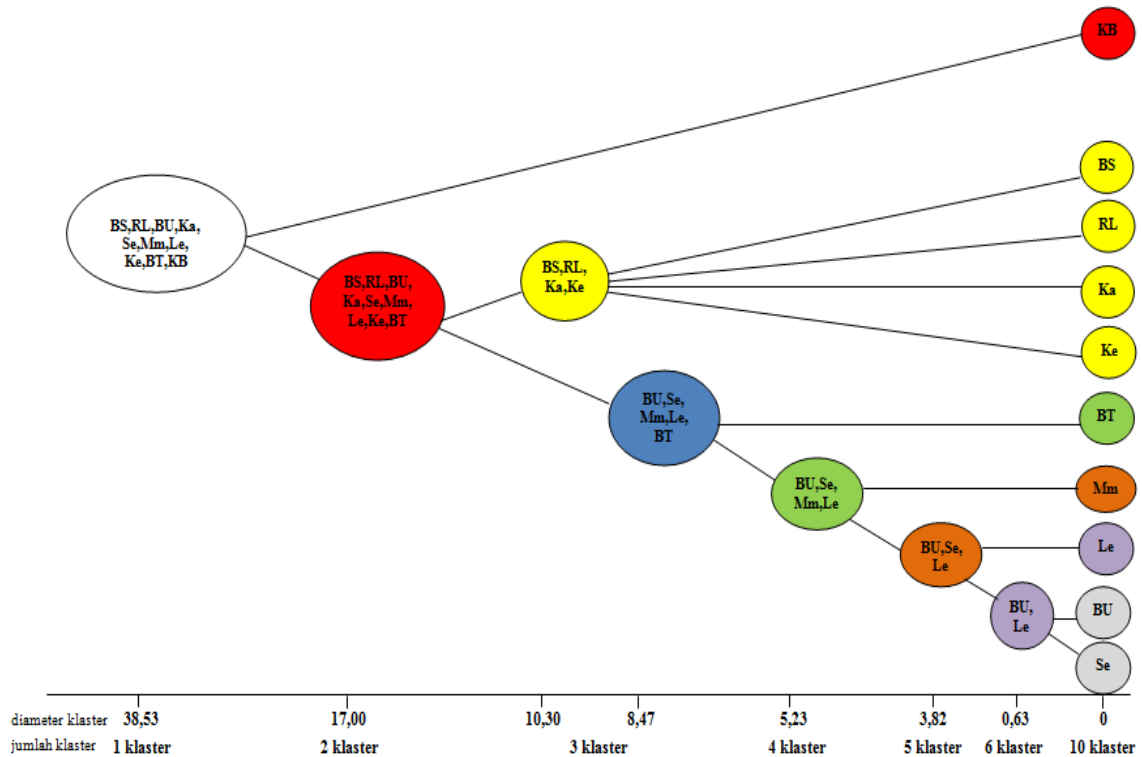
Berdasarkan Tabel 6 proses pengklasteran hierarki divisif dimulai dari satu kluster utama. Langkah 1 terjadi pemisahan obyek *KB* dengan jarak pemisahan terjauh sebesar 32,49. Sehingga obyek *KB* membentuk *splinter group 1*.

Langkah 2 adalah mengukur kedekatan antar obyek dengan cara menghitung selisih diantara semua obyek sisaan dengan obyek yang terdapat dalam *splinter group 1*. Hasil selisih yang bernilai negatif menunjukkan bahwa tidak ada obyek sisaan yang memiliki jarak kedekatan dengan *splinter group 1*. Berdasarkan algoritma pengklasteran devisif, proses dilanjutkan dengan cara memilih jarak rata-rata ketaksamaan yang paling tinggi dari semua obyek sisaan. Sehingga diperoleh obyek *BS* yang membentuk *splinter group 2* dengan jarak rata-rata pemisahan sebesar 11,37.

Langkah 3 diperoleh obyek *RL* yang memiliki jarak kedekatan ke *splinter group 2* yaitu 1,86. Sehingga diperoleh jarak pemisahan obyek *RL* dari kluster utama sebesar 9,51. Selanjutnya pada langkah ke 4 dan ke 5 diperoleh obyek *Ka* yang bergabung ke dalam *splinter group 2* dengan jarak kedekatan sebesar 1,37. Sehingga jarak pemisahan dari obyek *Ka* terhadap kluster utama yaitu 9,07. Kemudian obyek *Ke* juga bergabung ke dalam *splinter group 2* dengan jarak kedekatan sebesar 0,08. Sehingga diperoleh nilai jarak pemisahan obyek *Ke* sebesar 8,99.

Langkah 6 tidak ada obyek yang memiliki jarak kedekatan dengan *splinter group 2*. Sehingga terbentuklah *splinter group 3* yang memuat obyek *BT* dengan jarak rata-rata pemisahan sebesar 8,24. Selanjutnya pada langkah 7 dan 8 terjadi pemisahan obyek sehingga terbentuklah *splinter group 4* yang memuat obyek *Mm* dengan jarak rata-rata pemisahan yaitu 4,83. Kemudian terbentuk *splinter group 5* yang memuat obyek *Se*

dengan jarak rata-rata pemisahan sebesar 3,61. Langkah 9 terbentuk *splinter group* 6 dengan anggota obyek  $\{BU, Le\}$  dengan jarak pemisahan sebesar 0,63. Pada langkah terakhir ini semua obyek harus dipisahkan menjadi 10 kluster. Bagan dari pengklasteran hierarki divisif dapat dilihat pada Gambar 1 berikut.



**Gambar 1. Bagan dari proses pengklasteran hierarki divisif**

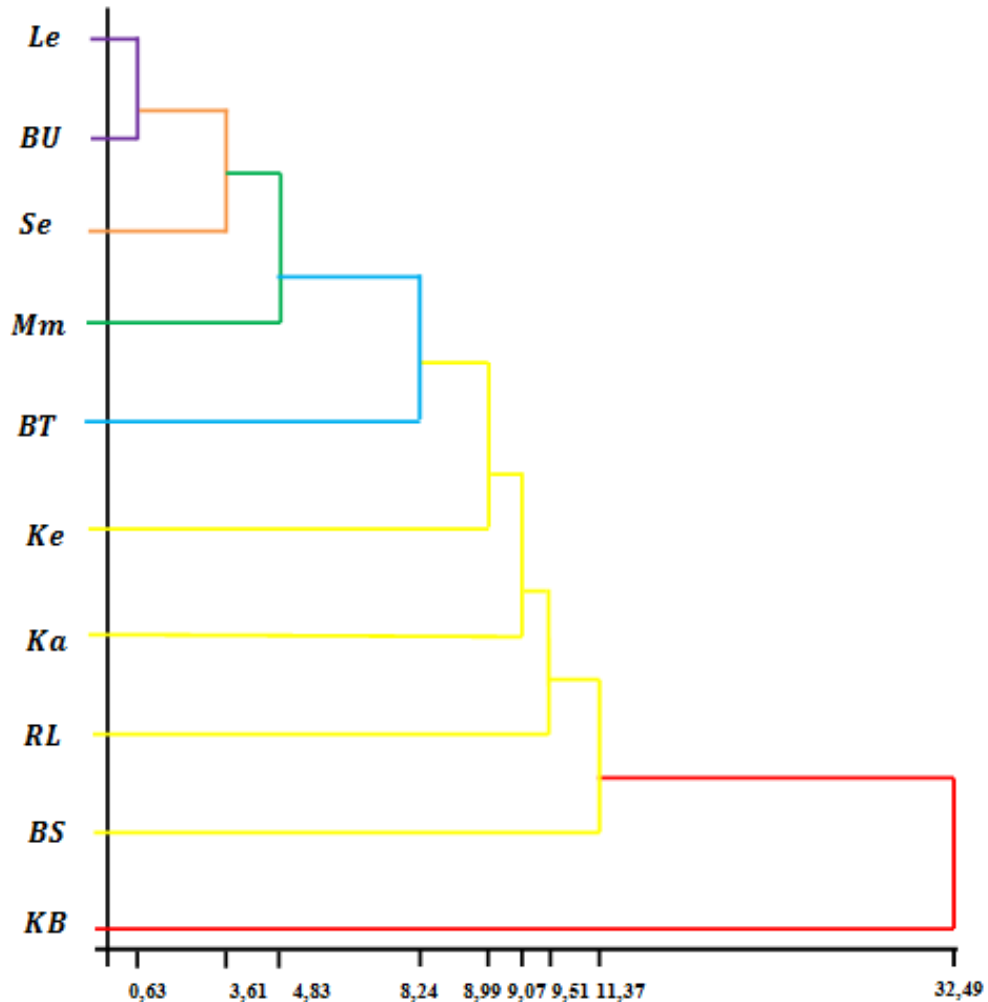
Berdasarkan Gambar 1 diperoleh sejumlah kluster seperti pada Tabel 7.

**Tabel 7. Jumlah kluster dari analisis kluster hierarki divisif**

Langkah	Jumlah kluster	Anggota kluster
0	1	$\{BS, RL, BU, Ka, Se, Mm, Le, Ke, BT, KB\}$
1	2	$\{BS, RL, BU, Ka, Se, Mm, Le, Ke, BT\}, \{KB\}$
2	3	$\{KB\}, \{RL, BU, Ka, Se, Mm, Le, Ke, BT\}, \{BS\}$
3	3	$\{KB\}, \{BU, Ka, Se, Mm, Le, Ke, BT\}, \{BS, RL\}$
4	3	$\{KB\}, \{BU, Se, Mm, Le, Ke, BT\}, \{BS, RL, Ka\}$
5	3	$\{KB\}, \{BU, Se, Mm, Le, BT\}, \{BS, RL, Ka, Ke\}$
6	4	$\{KB\}, \{BS, RL, Ka, Ke\}, \{BU, Se, Mm, Le\}, \{BT\}$
7	5	$\{KB\}, \{BS, RL, Ka, Ke\}, \{BT\}, \{BU, Se, Le\}, \{Mm\}$
8	6	$\{KB\}, \{BS, RL, Ka, Ke\}, \{BT\}, \{Mm\}, \{BU, Le\}, \{Se\}$
9	10	$\{KB\}, \{BS\}, \{RL\}, \{Ka\}, \{Ke\}, \{BT\}, \{Mm\}, \{BU\}, \{Le\}, \{Se\}$

## 6. Dendrogram dan Jumlah Klaster

Hasil dari algoritma pengklasteran hierarki divisif dapat digambarkan ke dalam dendrogram seperti pada Gambar 2.



Gambar 2. Dendrogram dari hasil analisis klaster hierarki dengan jarak *Euclid*

Setelah diperoleh beberapa *splinter group* dan dendrogram pada Gambar 2, penulis dapat mengambil 2 pemisahan klaster.

Tabel 8. Dua pemisahan klaster

Klaster	Anggota Klaster	Jarak pemisahan	Jarak rata-rata klaster
klaster 1	<i>KB</i>	32,49	32,49
klaster 2	<i>BS</i>	11,37	6,32
	<i>RL</i>	9,51	
	<i>Ka</i>	9,07	
	<i>Ke</i>	8,99	
	<i>BT</i>	8,24	
	<i>Mm</i>	4,83	
	<i>Se</i>	3,61	
	<i>BU</i>	0,63	
	<i>Le</i>	0,63	

Berdasarkan Tabel 8 diperoleh kluster 1 yang memuat Kota Bengkulu. Nilai jarak rata-rata pemisahan dari kluster 1 ini menunjukkan tingkat pendidikan tertinggi penduduk yang berumur 15 tahun ke atas di Kota Bengkulu sebesar 32,49. Sedangkan untuk kluster 2 terdiri dari 9 kabupaten yaitu Bengkulu Selatan, Rejang Lebong, Kaur, Kepahiang, Bengkulu Tengah, Mukomuko, Seluma, Bengkulu Utara dan Lebong. Nilai jarak rata-rata pemisahan dari 9 kabupaten menunjukkan tingkat pendidikan tertinggi penduduk yang berumur 15 tahun ke atas pada kluster 2 sebesar 6,32. Hasil dari 3 pemisahan kluster adalah sebagai berikut.

**Tabel 9. Tiga pemisahan kluster**

<b>Klaster</b>	<b>Anggota Klaster</b>	<b>Jarak pemisahan</b>	<b>Jarak klaster</b>
klaster 1	<i>KB</i>	32,49	32,49
klaster 2	<i>BS</i>	11,37	9,375
	<i>RL</i>	1,86	
	<i>Ka</i>	1,37	
klaster 3	<i>BT</i>	7,96	3,59
	<i>Ke</i>	6,29	
	<i>Mm</i>	4,83	
	<i>Se</i>	3,61	
	<i>BU</i>	0,63	
	<i>Le</i>	0,63	

Diketahui dari Tabel 9 Kota Bengkulu termasuk dalam kluster 1. Nilai jarak rata-rata pemisahan dari Kota Bengkulu menunjukkan tingkat pendidikan tertinggi penduduk yang berumur 15 tahun ke atas sebesar 32,49. Pada kluster 2 terdapat 4 kabupaten yaitu Bengkulu Selatan, Rejang Lebong, Kaur dan Kepahiang. Berdasarkan jarak rata-rata pemisahan dari 4 kabupaten tersebut diketahui tingkat pendidikan tertinggi penduduk yang berumur 15 tahun ke atas pada kluster 2 sebesar 9,735. Sedangkan kluster 3 terdiri dari 5 kabupaten yaitu Bengkulu Tengah, Mukomuko, Seluma, Bengkulu Utara dan Lebong. Berdasarkan jarak rata-rata pemisahan dari 5 kabupaten tersebut diketahui tingkat pendidikan tertinggi penduduk yang berumur 15 tahun ke atas sebesar 3,59.

#### **D. KESIMPULAN DAN SARAN**

##### **1. Kesimpulan**

Hasil algoritma pengklasteran hierarki divisif dari data penelitian diperoleh 6 *splinter group*. *Splinter group* tersebut digambarkan ke dalam sebuah dendogram. Berdasarkan ukuran kemiripan yang terdapat dalam dendogram tersebut, diambil 3 pemisahan kluster seperti berikut:

- a) Kluster 1 yang memuat Kota Bengkulu dengan jarak pemisahan sebesar 32,49. Berdasarkan nilai jarak pemisahan tersebut penulis menamakan kluster 1 sebagai

kota dengan tingkat pendidikan penduduk yang baik, dimana Kota Bengkulu memiliki jumlah penduduk dengan tingkat pendidikan menengah dan pendidikan tinggi terbanyak.

- b) Klaster 2 yang terdiri dari Bengkulu Selatan, Rejang Lebong, Kaur dan Kepahiang dengan jarak pemisahan sebesar 9,735. Berdasarkan nilai jarak pemisahan tersebut penulis menamakan klaster 2 sebagai kabupaten dengan tingkat pendidikan penduduk yang cukup baik, dimana 4 kabupaten tersebut memiliki jumlah penduduk dengan pendidikan dasar dan pendidikan menengah yang cukup banyak.
- c) Klaster 3 yang terdiri dari Bengkulu Tengah, Mukomuko, Seluma, Bengkulu Utara dan Lebong dengan jarak pemisahan sebesar 3,59. Berdasarkan nilai jarak pemisahan tersebut penulis menamakan klaster 3 sebagai kabupaten dengan tingkat pendidikan penduduk yang kurang baik, dimana 5 kabupaten tersebut memiliki jumlah penduduk yang tidak sekolah dan pendidikan dasar terbanyak.

## 2. Saran

Penelitian mengenai analisis klaster hierarki divisif dapat diaplikasikan pada berbagai bidang ilmu lainnya. Penelitian lanjutan yang disarankan adalah membandingkan hasil analisis klaster dari beberapa metode dalam analisis klaster.

## DAFTAR PUSTAKA

- [1] BPS. 2016. *Provinsi Bengkulu dalam Angka 2016*. BPS Provinsi Bengkulu. Bengkulu.
- [2] Chalik, I. 2016. Tingkat Pendidikan di Bengkulu Masih Rendah. <http://www.tingkat.pendidikan.di.bengkulu.masih.rendah/ANTARA.news.terkini.htm> diakses tanggal 11 Februari 2017.
- [3] Jobson, J.D. 1992. *Applied Multivariate Data Analysis*. Springer-Verlag Inc. New York
- [4] Kaufman, L and Rousseeuw, P.J. 2005. *Finding Groups in Data*. 9<sup>nd</sup> Edition. John Wiley & Sons. Ney Jersey.
- [5] Nugroho, S. 2008. *Statistika Multivariat Terapan*. Edisi Pertama. UNIB PRESS. Bengkulu
- [6] Sharma, S. 1996. *Applied Multivariate Technique*. John Wiley & Sons. USA.

# ANALISIS KLASTER DENGAN MENGGUNAKAN METODE *SINGLE LINKAGE* DAN METODE *K-MEANS*

(Studi Kasus: Data Produksi, Produktivitas, dan Luas Panen Tanaman Sayuran di  
Provinsi Bengkulu Tahun 2013)

Rezki Wahyuni<sup>1</sup>, Sigit Nugroho<sup>2</sup>, dan Pepi Novianti<sup>2</sup>

<sup>1</sup>Mahasiswa Jurusan Statistika, Prodi Matematika, FMIPA, UNIB, Bengkulu

<sup>2</sup>Dosen Pembimbing Jurusan Statistika, Prodi Matematika, FMIPA, UNIB, Bengkulu

[Rizqhywahyuni@gmail.com](mailto:Rizqhywahyuni@gmail.com)

## ABSTRAK

Analisis kluster merupakan teknik dalam analisis multivariat yang mempunyai tujuan untuk mengelompokkan obyek-obyek pengamatan menjadi beberapa kelompok berdasarkan karakteristik yang dimilikinya. Analisis kluster terdiri dari kluster hierarki dan kluster non hierarki. Salah satu metode pada kluster hierarki adalah kluster *Single Linkage*, dan kluster non hierarki adalah *K-Means*. Penelitian ini bertujuan untuk mengetahui hasil pengklasteran di Provinsi Bengkulu per Kabupaten, dan untuk mengetahui metode yang baik dari hasil pengklasteran metode *Single Linkage* dan metode *K-Means*, berdasarkan Data Produksi, Produktivitas, dan Luas Panen Tanaman Sayuran di Provinsi Bengkulu Tahun 2013. Penelitian ini menggunakan jarak *Euclid*. Berdasarkan hasil penelitian diperoleh pengelompokan data menjadi tiga anggota kluster yang terbentuk, dan memiliki hasil yang berbeda-beda dari setiap data dan metode yang digunakan. Dari hasil pengklasteran menyatakan bahwa metode *Single Linkage* merupakan metode yang baik dan Kabupaten Rejang Lebong memiliki pengaruh yang baik dengan variabel Cabe.

**Kata Kunci:** Analisis Kluster, Metode *Single linkage*, Metode *K-Means*, Jarak *Euclid*

## PENDAHULUAN

Indonesia merupakan negara agraris yang memiliki sumber daya alam yang beraneka ragam dan memiliki wilayah yang cukup luas. Letak astronomis negara Indonesia  $6^{\circ}$  LU –  $11^{\circ}$  LS dan  $94^{\circ}$  BT –  $141^{\circ}$  BT menandakan bahwa wilayah Indonesia merupakan wilayah yang subur dan beriklim tropis. Iklim di Indonesia yang cukup dalam memperoleh sinar matahari sepanjang tahun, mempengaruhi tumbuh suburnya setiap tanaman dengan mudah. Potensi wilayah yang demikian sangat baik kaitannya dalam pengembangan sektor pertanian (Anonim, 2009).

Sektor pertanian berperan penting pula dalam perekonomian Provinsi Bengkulu karena merupakan sektor utama yang memberikan peranan terbesar dalam pembentukan Produk Domestik Regional Bruto (PDRB). Pada tahun 2013 peranan sektor pertanian terhadap PDRB Provinsi Bengkulu adalah 39,84 persen dengan nilai nominal 8,4 triliun rupiah (atas dasar harga berlaku). Sektor pertanian merupakan kegiatan ekonomi yang menonjol untuk daerah ini bila dibandingkan dengan kegiatan di sektor lain-lainnya (Hamidi, 2013).

Salah satu kegiatan pertanian yang memiliki potensi untuk pertumbuhan baru adalah komoditas hortikultura. Hortikultura merupakan suatu cabang dari ilmu pertanian yang mempelajari budidaya buah-buahan, sayuran dan tanaman hias. Provinsi Bengkulu juga memiliki keragaman produksi tanaman hortikultura seperti sayuran dan buah-buahan. Sayuran merupakan komoditas pertanian yang memiliki harga cukup tinggi di pasaran. Tanamansayuran dapat secara nyata mendatangkan keuntungan bagi petani, sehingga banyak petani Indonesia yang membudidayakan tanaman sayuran (Hernanto, 2005).

Terdapat 8 jenis komoditi sayuran yang tercatat ditanam di berbagai daerah kabupaten/kota di Provinsi Bengkulu, yaitu bawang merah, cabe, bawang daun, tomat, terong, ketimun, kangkung dan bayam. Total luas panen tanaman sayuran pada tahun ini mencapai 997.153 hektar dengan produksi sebesar 2.251,635 ton (Hamidi, 2013). Namun, produktivitas tanaman hortikultura khususnya sayuran pada setiap kecamatan tentunya memiliki karakteristik yang berbeda-beda, melihat luas areal tanaman menghasilkan, iklim/cuaca, kesuburan tanah, dll. Melihat kondisi tersebut, maka perlu adanya pengelompokan komoditi dan kabupaten berdasarkan luas areal dan produksi yang berguna untuk meningkatkan produktivitas tanaman hortikultura di Provinsi Bengkulu (Anonim, 2009).

Analisis kluster sering kali ditemui di kehidupan sehari-hari, baik yang terkait dengan bidang sosial, bidang kesehatan, bidang *marketing*, maupun bidang pertanian. Analisis kluster merupakan suatu teknik yang dapat digunakan untuk mengklasifikasikan obyek atau kasus kedalam kelompok yang relatif homogen. Obyek atau kasus dalam setiap kelompok cenderung mirip satu sama lain dan berbeda jauh (tidak sama) dengan objek dari kluster lainnya. Analisis kluster juga disebut analisis klasifikasi atau taksonomi numerik (Oktarina, 2014).

Ditinjau dari hal-hal yang dikelompokkan analisis kluster dibagi menjadi dua macam yaitu pengelompokan observasi dan pengelompokan variabel. Sedangkan secara umum analisis kluster dibagi menjadi dua metode yaitu metode hierarki dan metode non hierarki. Di dalam metode hierarki sendiri terdapat beberapa metode, metode-metode yang termasuk dalam metode hierarki adalah metode Pautan Tunggal (*Single Linkage*), metode Pautan Lengkap (*Complete Linkage*), metode Antar Pusat (*Centroid Linkage*), metode Pautan Rata-rata (*Average Linkage*) dan Metode Ward (*Ward's Method*), sedangkan metode yang termasuk dalam metode non hirarki adalah metode *K-Means* (Soraya, 2011).

Metode dalam analisis kluster mempunyai kelemahan dan kelebihan sendiri - sendiri. Diantaranya adalah jika ditinjau dari proses pengelompokannya, pada metode hierarki proses pengelompokannya terjadi secara alami sehingga kelompoknya terbentuk dengan sendirinya dan digambarkan dalam bentuk *Dendogram*, sedangkan pada metode non hierarki jumlah kelompok yang terbentuk harus ditentukan terlebih dahulu oleh peneliti.

Berdasarkan uraian diatas, penulis tertarik untuk meneliti tentang Analisis Kluster dengan Menggunakan Metode *Single Linkage* dan Metode *K-Means*. Guna mengelompokkan komoditi dan kabupaten berdasarkan produksi, produktivitas, dan luas panen tanaman sayuran, agar dapat meningkatkan produktivitas tanaman hortikultura di Provinsi Bengkulu.

## Tinjauan Pustaka

### Analisis Kluster

Analisis kluster adalah suatu metode statistik yang mengidentifikasi kelompok sampel berdasarkan karakteristik serupa. Analisis kluster mengelompokkan elemen mirip sebagai objek penelitian yang mempunyai tingkat homogenitas yang tinggi antar objek menjadi kluster yang berbeda dengan tingkat heterogenitas obyek yang tinggi antar kluster. Pengklasteran ini didasarkan pada gugus variabel yang dipertimbangkan untuk diteliti (Nugroho, 2008).

Analisis kluster merupakan suatu metode yang membagi individu kedalam kelompok yang bermakna. Analisis kluster mengelompokkan obyek berdasarkan informasi yang diperoleh pada data yang menggambarkan obyek dan keterkaitannya (Johnson and Wichern, 2002).

Analisis kluster harus memenuhi dua asumsi yaitu:

1. Sampel yang diambil harus benar-benar dapat mewakili populasi.
2. Tidak ada multikolinieritas (korelasi antar obyek), jika ada maka nilai multikolinieritasnya kurang dari 0,5 (Nugroho, 2008).

### Proses Analisis Kluster

#### 1. Standarisasi atau Transformasi Variabel

Standarisasi variabel dilakukan apabila terdapat perbedaan satuan nilai yang besar antar variabel. Misalnya ada yang dalam satuan juta dan ada yang satuan puluhan bahkan lebih kecil. Perbedaan data yang besar akan menyebabkan perhitungan jarak menjadi tidak valid sehingga data harus ditransformasi.

Transformasi dapat dilakukan dengan uji *z-score* (Nugroho, 2008), yaitu transformasi data dalam bentuk normal baku  $N(0,1)$  yang dirumuskan sebagai berikut :

$$z_i = \frac{x_i - \bar{x}}{s} \quad \text{untuk } i = 1, 2, \dots, n \quad (1.1)$$

Keterangan :

$x_i$  = data ke  $i$

$\bar{x}$  = rata - rata data

$s$  = simpangan baku

#### 2. Pengukuran Konsep Kemiripan Jarak antar Objek

Metode ini berbentuk matrix simetri  $n \times n$  yang berisi kemiripan atau ketakmiripan. Konsep ketakmiripan pada metode ini berdasarkan pada ukuran jarak antar objek, di mana jarak yang besar menunjukkan sedikit kesamaan dan jarak yang pendek atau kecil menunjukkan bahwa suatu objek semakin mirip dengan objek yang lainnya. Fungsi jarak objek  $a$  dengan objek  $b$  dapat dinotasikan dengan  $d(a, b)$  yang memenuhi sifat-sifat berikut (Nugroho, 2008).

1.  $d(a, b) \geq 0$
2.  $d(a, a) = 0$
3.  $d(a, b) = d(b, a)$
4.  $(a, b)$  meningkat seiring tidak miripnya  $a$  dan  $b$
5.  $d(a, c) \leq d(a, b) + d(b, c)$

Menurut Nugroho (2008), pengukuran jarak ada bermacam-macam, pada penelitian ini jarak yang akan digunakan adalah jarak *Euclid*.

Jarak *Euclidean* adalah besarnya jarak suatu garis lurus yang menghubungkan antar obyek. Jarak *Euclid* merupakan akar jumlah kuadrat perbedaan nilai untuk tiap variabel. Jarak *Euclid* dirumuskan sebagai berikut (Nugroho, 2008).

Jika  $\mathbf{x}' = (x_1, x_2, \dots, x_p)$  dan  $\mathbf{y}' = (y_1, y_2, \dots, y_p)$  maka  $d(x, y)$  adalah jarak antara objek  $x$  dengan objek  $y$  (Johnson and Wichern, 2002).

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_p - y_p)^2}$$

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(\mathbf{x} - \mathbf{y})'(\mathbf{x} - \mathbf{y})} = \sqrt{\sum_{i=1}^p (x_i - y_i)^2} \quad (1.2)$$

Keterangan :

$d(\mathbf{x}, \mathbf{y})$	= jarak antara obyek $x$ dan $y$
$x_i$	= nilai obyek $x$ pada variabel $i$
$y_i$	= nilai obyek $y$ pada variabel $i$
$p$	= banyak variabel yang diamati

## Metode Pengklasteran

### 1. Metode Hierarki

Metode kluster hierarki (*hierarchical clustering methods*) adalah metode yang digunakan untuk mengelompokkan suatu pengamatan secara terstruktur berdasarkan kemiripan sifat dan banyaknya kelompok yang dapat dibentuk belum diketahui. Untuk penyajian banyaknya kelompok yang dapat dibentuk yaitu dari kelompok terbesar hingga terkecil dan sebaliknya. Pengklasteran hierarki dimulai dengan satu kluster yang diperoleh dari hasil pengamatan suatu objek dan diakhiri dengan kluster yang berisi semua objek atau sebaliknya (Rencher, 2002).

### Metode Penggabungan (*Agglomerative Methods*)

Metode penggabungan (*agglomerative methods*) merupakan metode pengklasteran yang dilakukan dari pengamatan objek-objek yang memiliki kesamaan sehingga bergabung menjadi suatu kluster kecil dan dari kluster-kluster kecil yang terbentuk akan digabungkan menjadi satu kluster besar yang berisikan semua kluster (Rencher, 2002).

### *Single Linkage (Pautan Tunggal)*

Metode pautan tunggal adalah proses pengklasteran yang didasarkan pada jarak terdekat antar objek. Jika dua objek terpisah oleh jarak yang pendek, maka kedua objek tersebut akan digabung menjadi satu kluster. Metode ini dimulai dengan menentukan objek-objek yang memiliki jarak terkecil dalam matriks *proximity*  $\mathbf{D} = \{d_{ik}\}$ . Objek-objek yang memiliki jarak terkecil digabungkan dalam satu kluster, misal  $U$  dan  $V$  adalah dua objek yang akan diklasterkan sehingga diperoleh kluster  $(UV)$ . Untuk mencari jarak antara kluster  $(UV)$  dan kluster  $W$  atau kluster lainnya diperoleh dengan rumus:

$$d_{(UV)W} = \min\{d_{UW}, d_{VW}\} \quad (1.3)$$

Keterangan:

$d_{UW}$	adalah jarak terdekat dari kluster $U$ dan $W$
$d_{VW}$	adalah jarak terdekat dari kluster $V$ dan $W$

### 2. Analisis Kluster Non Hierarki

Analisis kluster dengan metode non hierarki merupakan metode kluster yang menentukan jumlah kluster secara manual (Baroroh, 2012). Teknik analisis kluster nonhierarki dirancang untuk mengelompokkan item bukan variabel menjadi kumpulan  $K$  kluster. Banyaknya kluster  $K$ , ditentukan terlebih dahulu untuk memulai prosedur pengklasteran. Metode ini dapat diterapkan pada himpunan data yang jauh lebih besar dibandingkan menggunakan teknik analisis kluster hierarki.

Metode analisis kluster non hierarki berhubungan dengan *K-Means*, ada tiga pendekatan yang digunakan untuk menempatkan masing-masing observasi dalam satu kluster yaitu:

- a. *Sequential Threshold*
- b. *Parallel Threshold*
- c. *Optimization*

### **K – Means**

*K-Means* merupakan metode analisis kluster non hierarki yang dimulai dengan menentukan terlebih dahulu jumlah kluster yang diinginkan. Setelah jumlah kluster diketahui, baru proses kluster dilakukan tanpa mengikuti proses hierarki (Santoso, 2010). *K-Means* banyak digunakan di berbagai bidang karena sederhana dan mudah diimplementasikan. Metode ini mengklusterkan secara *partitioning* yang memisahkan data ke kelompok yang berbeda. Pembagian data biasanya dibentuk ke dalam 2 atau lebih kelompok data (Supranto, 2004).

Penggunaan kluster *K-Means* untuk menjelaskan algoritma dalam penentuan suatu obyek ke dalam kluster tertentu berdasarkan rataan terdekat. Asumsikan  $n$  adalah obyek dan  $p$  adalah variabel, dinotasikan  $x(i, j)$  adalah data obyek ke  $i$  dan variabel ke  $j$ , dengan  $i = 1, 2, \dots, n$  dan  $j = 1, 2, \dots, p$  dengan menggunakan jarak *euclid* antar objek.

Jika  $p(n, k)$  adalah partisi yang merupakan hasil pada tiap objek dialokasikan untuk salah satu dari kluster ke  $1, 2, \dots, k$ . Rata-rata variabel ke  $j$  pada kluster ke  $l$  dinotasikan dengan  $\bar{x}(j, l)$ . Maka jarak antara objek ke  $i$  dan kluster ke  $l$  didefinisikan sebagai berikut:

$$d(i, l) = \left( \sum_{j=1}^p [x(i, j) - \bar{x}(j, l)]^2 \right)^{1/2} \quad (1.4)$$

dan error partisi:

$$E[p(n, k)] = \sum_{i=1}^n D[i, l(i)]^2 \quad (1.5)$$

$E[p(n, k)]$  adalah error partisi,  $l(i)$  adalah kluster yang memuat obyek ke  $i$ ,  $D[i, l(i)]$  adalah jarak *euclid* antara obyek ke  $i$  dan rata-rata kluster yang memuat obyek (Nugroho, 2008).

### **Interpretasi Profil**

Interpretasi profil meliputi pengujian tiap kluster untuk menamai dan menandai dengan suatu label yang secara akurat dapat menjelaskan kealamian kluster. Tahap ini menggambarkan karakteristik tiap kluster untuk menjelaskan kluster-kluster tersebut dapat berbeda pada dimensi yang relevan. Membuat interpretasi dan profil kluster tidak hanya untuk memperoleh suatu gambaran saja melainkan menyediakan suatu rata-rata untuk menilai korespondensi pada kluster yang telah terbentuk dan profil kluster memberikan arahan bagi penilaian terhadap signifikansi (Nugroho, 2008).

## **METODE PENELITIAN**

### **Data**

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) yang telah dipublikasikan dalam bentuk buku dengan judul Sektor Pemerintah dalam Pertanian Indonesia. Data yang digunakan berupa data produksi tanaman sayuran, produktivitas tanaman sayuran, dan luas panen tanaman sayuran di Provinsi Bengkulu perKabupaten pada Tahun 2013.

### **Tahapan Analisis Data**

Tahapan – tahapan yang dilakukan untuk menganalisis data adalah sebagai berikut:

1. Mempartisi data untuk dikelompokan
2. Melakukan uji multikolinieritas dengan menggunakan korelasi dan *Variance Inflation Factor* (VIF)
3. Memilih ukuran jarak antar objek dengan menggunakan jarak euclid dan mengolah data menggunakan analisis kluster hierarki yaitu metode *Single Linkage* dan non hierarki yaitu metode *K-Means* dengan bantuan program R 2.14.1.
4. Melakukan interpretasi pada kluster yang telah di peroleh.

### **Hasil dan Pembahasan**

#### **Deskripsi Data**

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) di Provinsi Bengkulu Tahun 2013. Data yang digunakan terdiri dari 3 data yang berbeda yaitu, data produksi, produktivitas, dan luas panen tanaman sayuran. Variabel yang digunakan dalam penelitian ini ada 8 variabel yang terdiri dari, bawang merah, cabe, bawang daun, tomat, terung, ketimun,

kangkung, dan bayam. Sedangkan obyek-obyek yang digunakan adalah 10 Kabupaten yang berada di Provinsi Bengkulu.

### Uji Multikolinieritas

Multikolinieritas terjadi bila terdapat dua atau lebih variabel yang saling berkaitan. Untuk mengetahui ada tidaknya multikolinieritas antar variabel dapat dilakukan dengan matrik korelasi. Pada penelitian ini untuk mengetahui ada tidaknya multikolinieritas, penulis menggunakan uji korelasi dan uji *Variance Inflation Factor* (VIF) untuk mendeteksi masalah multikolinieritas. Jika nilai korelasi mendekati satu atau lebih besar dari 0,5 maka dapat dikatakan terdapat masalah multikolinieritas.

Nilai matrix korelasi dari output program R dapat dilihat pada lampiran 3, dari matrix korelasi pada lampiran 3 terdapat nilai korelasi yang lebih dari 0,5 dan terdapat matrix korelasi yang nilainya kurang dari 0,5. Variabel yang memiliki nilai matrix korelasi kurang dari 0,5 dapat dilihat pada tabel 1 berikut:

Tabel 4 Matrix Korelasi Data Produksi Tanaman Sayuran Tahun 2013

	Cabe	Kangkung	Ketimun	Terung	Tomat
B. Daun	0,0974736	0,396984	0,41681823	0,0238983	
B. Merah	0,007417523	0,313822489	0,321295434	0,081170266	
Bayam	0,1765056		0,4599876	0,1118327	
Cabe					0,2226143
Terung					0,14424849

Berdasarkan data produksi tanaman sayuran maka didapatkan nilai *Variance Inflation Factor* (VIF) yang diperoleh dari output program R dan mendapatkan hasil sebagai berikut.

Tabel 2 Nilai *Variance Inflation Factor* (VIF) Data Produksi Tanaman Sayuran dari Output Program R

No	Variabel	VIF
1	Bawang Merah	6,241815
2	Ketimun	1,697112
3	Kangkung	2,376218
4	Bayam	8,140464

### Metode Pengklasteran

Pengklasteran pada penelitian ini menggunakan metode Hierarki dan metode Non Hierarki, pada hierarki metode yang digunakan yaitu metode *Single Linkage*, sedangkan pada non hierarki metode yang digunakan yaitu *K-Means*, jarak yang digunakan oleh kedua metode ini adalah Jarak *Euclid*.

### Metode Hierarki (*Single Linkage*)

Pengklasteran pada penelitian ini menggunakan metode Hierarki yaitu Metode *SingleLinkage* dan jarak yang digunakan adalah Jarak *Euclid*, tabel dibawah ini menunjukkan masing-masing nilai dari jarak *Euclid* yang dihasilkan oleh 3 data.

Tabel 3 Nilai Jarak *Euclid*, pada Output Dari Program R

	Produksi	Produktivitas	Luas Panen
<b>Min</b>	2145,4	236,63	48,888
<b>1st Qu</b>	4930,8	614,02	256,280
<b>Median</b>	6745,6	937,31	369,320
<b>Mean</b>	59197,0	1299,90	756,790
<b>3rd Qu</b>	73069,0	1746,00	536,030
<b>Max</b>	216430,0	3240,60	2586,800

Berdasarkan Tabel 11 dapat dilihat nilai-nilai dari masing-masing jarak *Euclid* yang dihasilkan oleh 3 data. Dari jarak tersebut dapat dilihat bahwa terdapat perbedaan dari nilai minimum, kuartil pertama,

*median, mean, kuartil ketiga dan nilai maksimum pada masing-masing data yang dihasilkan oleh jarak Euclid.*

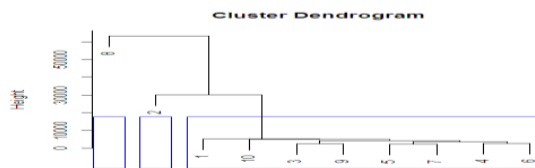
**Metode *Single Linkage* untuk Data Produksi Tanaman Sayuran**

Tabel 4 Proses Agglomerative metode *Single Linkage* dan nilai Jarak *Euclid* pada Output Dari Program R

Step	Clusters	Joined	Order of Object	Distance Level
1	5	7	3	5171,139
2	3	9	9	2462,161
3	4	6	4	4211,873
4	3	1	6	2662,936
5	2	4	5	3498,428
6	5	10	7	2145,444
7	1	6	10	4894,157
8	7	8	8	68470,311
9	8	2	2	204976,039

Setelah proses *Agglomerative* terbentuk maka pengklasteran akan dilakukan dengan menggabungkan semua kluster-kluster kecil yang sudah terbentuk menjadi satu kluster besar yang berisikan semua kluster. Jika kluster besar sudah terbentuk maka hasil pengklasteran akan muncul dalam bentuk dendogram.

Dendrogram merupakan visualisasi proses pengklasteran yang terjadi. Posisi garis pada skala dendrogram menunjukkan jarak pada kluster yang sudah tergabung. Dendrogram yang diperoleh pada penelitian ini adalah sebagai berikut:



Gambar 1. Dendrogram Metode *Single Linkage* menggunakan Jarak *Euclid*

Pada dendrogram metode *Single Linkage* tahap pertama objek (Bengkulu Utara dan Bengkulu Tengah), objek (Seluma dan Lebong), objek (Kaur dan Muko-Muko), membentuk satu kluster tersendiri, karena mempunyai panjang garis yang sama dan tergabung menjadi satu kesatuan. Demikian seterusnya sampai akhirnya objek akan bergabung menjadi beberapa kluster.

Pada metode *Single Linkage*, kluster yang dapat dibentuk adalah 1,2 dan 3 kluster, dimana jika kluster 1 yang diambil maka anggotanya adalah kluster dengan objek (Bengkulu Selatan), objek (Kota Bengkulu), objek (Bengkulu Utara dan Bengkulu Tengah), objek (Seluma dan Lebong) dan objek (Kaur dan Muko-Muko). Jika kluster 2 yang diambil maka anggota kluster tersebut adalah objek (Rejang Lebong) dan 9 objek lainnya. Jika kluster 3 yang diambil maka anggotanya adalah kluster dengan objek (Kepahiang), dan 8 objek lainnya. Pada metode *Single Linkage* ini, Rejang Lebong membentuk kluster tunggal dengan anggota dirinya sendiri, dan Kepahiang juga membentuk kluster tunggal dengan anggota dirinya sendiri. Hal ini berarti objek Rejang Lebong dan Kepahiang berbeda jauh dengan objek pada kluster lainnya. Hal ini dikarenakan Rejang Lebong memiliki Hasil Produksi cabe dan daun bawang terbesar dari Provinsi lainnya. Dan objek Kepahiang juga memiliki Hasil Produksi cabe terbesar dari Provinsi lain.

**Metode Non Hierarki (*K-Means*)**

Metode *K-Means* memproses semua data secara sekaligus dimana *k* adalah banyaknya kluster. Dengan menggunakan R 2.14.1 maka data akan diproses.

**1. Metode *K-Means* untuk Data Produksi Tanaman Sayuran**

Berikut proses yang dilakukan dalam analisis klster metode *K-means* meliputi:

- a. Menentukan nilai *K* sebagai jumlah klster yang akan dibentuk

Dari output program R untuk proses pengklasteran dengan menggunakan metode *K-means* kluster yang akan dibentuk yaitu 3 kluster dengan memperhatikan semua kemungkinan pada data.

b. Menentukan nilai *centroid* awal secara *random*

Berdasarkan kluster yang akan dibentuk (*K*) pada proses pengklasteran metode *k-means* yang diambil secara acak. Sehingga terdapatlah nilai *centroid* awal metode *k-means* yang diperoleh/ditentukan dari tiga objek pengamatan yaitu, *centroid* pertama berupa Kabupaten Kepahiang, *centroid* kedua Kabupaten Rejang Lebong, *centroid* ketiga Kabupaten Kaur. Sehingga didapat nilai *centroid* awal sebagai berikut.

Tabel 5 Nilai *Centroid* awal metode *K-Means*

Variabel	Centroid Kluster		
	1	2	3
X <sub>1</sub>	0	692	0
X <sub>2</sub>	54646	9565	704
X <sub>3</sub>	20206	215140	305
X <sub>4</sub>	4554	16342	116
X <sub>5</sub>	44636	4136	1306
X <sub>6</sub>	20608	10325	175
X <sub>7</sub>	4756	3597	2572
X <sub>8</sub>	2306	7100	1213

c. Menentukan jarak setiap objek dengan setiap *centroid* pada metode *K-means*

Untuk menghitung jarak setiap objek dengan *centroid* pertama, kedua, dan ketiga, kita dapat menggunakan rumus jarak *euclid* pada persamaan (2.2). Adapun contoh perhitungan jarak pada setiap objek dengan *centroid* adalah sebagai berikut:

- Kabupaten Bengkulu Selatan dengan *Centroid* pertama

$$= \sqrt{(0 - 0)^2 + (4695 - 54646)^2 + (122 - 20206)^2 + \dots + (1661 - 2306)^2}$$

$$= 68467,27247$$

- Kabupaten Rejang Lebong dengan *centroid* pertama

$$= \sqrt{(692 - 0)^2 + (9568 - 54646)^2 + (21514 - 20206)^2 + \dots + (7100 - 2306)^2}$$

$$= 63205,30863$$

Lakukan cara yang sama untuk perhitungan jarak setiap objek dengan *centroid* pertama, kedua, dan ketiga, sampai objek ke 10. Sehingga diperoleh keseluruhan nilai jarak *euclid* dari masing-masing objek pada setiap *centroid* adalah sebagai berikut.

Tabel 6 Nilai jarak *Euclid* metode *K-means*

Objek	Kluster		
	1	2	3
Bengkulu Selatan	68467,27247	76272,7633	8241,167939
Rejang Lebong	63205,30863	0	31744,9797
Bengkulu Utara	71396,87443	75819,70775	5986,714124
Kaur	75009,63873	75989,5278	0
Seluma	72904,48184	76521,89155	3408,787321
Muko - Muko	74154,92218	75735,03059	2641,522856
Lebong	73032,81691	76880,78511	4755,261192
Kepahiang	0	92994,86998	75007,15603
Bengkulu Tengah	72563,90395	74606,1622	4527,419795
Kota Bengkulu	69770,10153	74621,09128	6138,120071

d. Mengelompokan objek berdasarkan jarak minimum

Misalkan angka 1 untuk nilai jarak yang kecil, sedangkan 0 untuk nilai jarak yang yang besar, maka dapat dilihat pada matriks berikut.

$$G = \begin{pmatrix} A & B & C & D & E & F & G & H & I & J \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \end{pmatrix}$$

Berdasarkan uraian diatas dapat dilihat anggota masing-masing kluster yang terbentuk, kluster pertama beranggotakan Kabupaten Kepahiang, kluster dua beranggotakan Kabupaten Rejang Lebong, sedangkan kluster ketiga beranggotakan delapan Kabupaten lainnya.

e. Iterasi 1, menentukan *centroid* baru

Berdasarkan hasil kluster yang terbentuk diatas, maka kita akan menentukan *centroid* baru, dengan cara yang sama pada langkah (b), maka diperoleh nilai *centroid* baru sebagai berikut.

Tabel 7 Nilai *Centroid* metode *K-Means*

Variabel	Centroid Kluster		
	1	2	3
X <sub>1</sub>	0	692	7
X <sub>2</sub>	54646	9565	3667
X <sub>3</sub>	20206	215140	90
X <sub>4</sub>	4554	16342	655
X <sub>5</sub>	44636	4136	2362
X <sub>6</sub>	20608	10325	1081
X <sub>7</sub>	4756	3597	1544
X <sub>8</sub>	2306	7100	1323

f. Iterasi 1, Menghitung jarak setiap objek dengan setiap *centroid* pada metode *K-means*

Lakukan cara yang sama pada langkah (c) untuk menghitung jarak setiap objek dengan *centroid* pertama, kedua, dan ketiga, kita dapat menggunakan rumus jarak *euclid* pada persamaan (2.2). Adapun contoh perhitungan jarak pada setiap objek dengan *centroid* adalah sebagai berikut:

- Kabupaten Bengkulu Selatan dengan *Centroid* pertama

$$= \sqrt{(0 - 0)^2 + (4695 - 54646)^2 + (122 - 20206)^2 + \dots + (1661 - 2306)^2}$$

$$= 68467,27247$$

- Kabupaten Rejang Lebong dengan *centroid* pertama

$$= \sqrt{(692 - 0)^2 + (9568 - 54646)^2 + (21514 - 20206)^2 + \dots + (7100 - 2306)^2}$$

$$= 63205,30863$$

Lakukan cara yang sama untuk perhitungan jarak setiap objek dengan *centroid* pertama, kedua, dan ketiga, sampai objek ke 10. Sehingga diperoleh keseluruhan nilai jarak *euclid* dari masing-masing objek pada setiap *centroid* adalah sebagai berikut.

Tabel 8 Nilai jarak *Euclid* metode *K-means*

Objek	Kluster		
	1	2	3
Bengkulu Selatan	68467,27247	76272,7633	5694,124779
Rejang Lebong	63205,30863	0	30549,31299
Bengkulu Utara	71396,87443	75819,70775	3899,311991
Kaur	75009,63873	75989,5278	3481,525528
Seluma	72904,48184	76521,89155	1632,812604
Muko - Muko	74154,92218	75735,03059	3389,776836
Lebong	73032,81691	76880,78511	3058,118866
Kepahiang	0	92994,86998	72036,22701
Bengkulu Tengah	72563,90395	74606,1622	2983,102412
Kota Bengkulu	69770,10153	74621,09128	3740,406128

g. Iterasi 1, Mengelompokan objek berdasarkan jarak minimum

Misalkan angka 1 untuk nilai jarak yang kecil, sedangkan 0 untuk nilai jarak yang yang besar, maka dapat dilihat pada matriks berikut.

$$G = \begin{matrix} & \begin{matrix} A & B & C & D & E & F & G & H & I & J \end{matrix} \\ \begin{matrix} A \\ B \\ C \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \end{pmatrix} \end{matrix}$$

Berdasarkan uraian diatas dapat dilihat anggota masing-masing kluster yang terbentuk, kluster pertama beranggotakan Kabupaten Kepahiang, kluster dua beranggotakan Kabupaten Rejang Lebong, sedangkan kluster ketiga beranggotakan delapan Kabupaten lainnya.

Setelah melakukan iterasi pertama ternyata tidak terdapat perpindahan dari masing-masing anggota kluster yang terbentuk, karena tidak terdapat perpindahan maka proses kluster dihentikan, untuk melihat hasil perhitungan pada output program R dapat dilihat pada lampiran 8 (a).

## Interpretasi

Interpretasi profil kluster adalah proses pemberian nama dari masing-masing kluster yang telah terbentuk. Nama yang diberikan sesuai keperluan peneliti, dengan syarat bahwa nama yang diberikan kepada kluster tersebut mewakili karakteristik atau ciri dari kelompok (Supranto, 2010).

## Kesimpulan

Dari hasil penelitian yang dilakukan dengan judul Analisis Kluster Dengan Menggunakan Metode *Single Linkage* dan Metode *K-Means* dengan menggunakan jarak *Euclid*, pengklusteran yang dilakukan dengan menggunakan tiga data yang berbeda yaitu Produksi, Produktivitas, dan Luas Panen Tanaman Sayuran ini menghasilkan pemecahan kluster yang berbeda antara metode *Single Linkage* dan metode *K-means*. Sehingga dapat disimpulkan sebagai berikut:

1. Dari hasil pengklusteran yang terbentuk dengan menggunakan metode *Single Linkage* dan metode *K-means*, berdasarkan Data Produksi, Produktivitas, dan Luas Panen Tanaman Sayuran, masing-masing metode membentuk 3 kluster yang berbeda dari setiap data tersebut. Pada metode *Single Linkage* kluster yang terbentuk tidak mendapatkan perbedaan yang jauh antara setiap data yang digunakan. Sedangkan metode *K-means* kluster yang terbentuk mendapatkan perbedaan yang jauh antar setiap data yang digunakan.
2. Pada penelitian ini dapat dijelaskan bahwa metode yang baik dalam analisis kluster yang terbentuk adalah metode *Single Linkage*, karena metode *Single Linkage* tidak mendapatkan perbedaan yang jauh antar setiap kluster yang terbentuk. Sedangkan Kabupaten yang paling berpengaruh pada penelitian ini adalah Kabupaten Rejang Lebong, karena selalu menghasilkan hasil yang baik dalam setiap proses kluster yang telah dibentuk.

## Saran

Pada tulisan ini penulis menggunakan metode *Single Linkage* dan *K-Means* dalam bidang ilmu pertanian. Penelitian selanjutnya dapat juga menggunakan data dalam bidang ilmu lain dengan menggunakan perbandingan metode *Single Linkage* dan *K-Means* dengan metode pengklusteran lain mengingat cakupan kluster yang sangat luas.

## DAFTAR PUSTAKA

- Anonim. 2009. *Pertanian Indonesia*. Penebar Swadaya. Jakarta.
- Hamidi, H. A. 2013. *Jurnal Optimis Pola Tanaman Usaha Tani Sayuran Selada dan Sawi di Daerah Produksi Padi*. Fakultas Pertanian Universitas Syiah Kuala. Banda Aceh.
- Hernanto, F. 2005. *Ilmu Usaha Tani*. Penebar Swadaya. Jakarta.  
<https://Pengertian.Produktivitas.kerja.htm>(Diakses 01 februari 2015).
- Johnson, R. A.,and D, W. Wichern. 2002. *Applied Multivariate Statistical Analysis, Fifth Edition*. Pearson Education. United States Of America.
- Nugroho, S. 2008. *Statistika Multivariat terapan, edisi pertama*. UNIB Press, Universitas Bengkulu.
- Oktarina, 2014. *Metode Linkage pada Persentase Kelahiran Balita Menurut Penolong Kelahirn Terakhir*. Skripsi, Program Studi Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam. Universitas Bengkulu.
- Rencher, A. C. 2002. *Methods Of Multivariate Analysis Second Edition*. Wiley Series In Probability and Mathematical Statistics: Canada.
- Soraya, Y. 2011. *Perbandingan Kinerja Metode Single Linkage Metode Complate Linkage Metode K-Means Dalam Analisis Kluster*. Jurusan Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Semarang.
- Supranto, J. 2004. *Analisis Multivariat Arti dan Interpretasi*. Rineka Cipta: Jakarta.

## **Analisis Konjoin Preferensi Mahasiswa Dalam Memilih Produk *Smartphone Android***

Irma Ismayasari<sup>1</sup>, Sigit Nugroho<sup>2</sup>, dan Etis Sunandi<sup>2</sup>

<sup>1</sup>Alumni Jurusan Statistika FMIPA Universitas Bengkulu

<sup>2</sup>Staf Pengajar Jurusan Statistika FMIPA Universitas Bengkulu

Abstrak pada Penelitian ini bertujuan untuk mengetahui preferensi dan kepentingan atribut terhadap spesifikasi produk *smartphone Android*. Hasil dari penelitian ini terlihat dimana responden lebih memilih *smartphone Android* dengan harga lebih terjangkau dan dengan ruang penyimpanan *memory* yang tidak terlalu besar. Performa atau daya tahan baterai yang tahan lama menjadi atribut dari pilihan responden karena dapat menunjang aktifitas pengguna *smartphone* dalam kebutuhan kesehariannya dan juga didukung dengan kekuatan jaringan yang cepat dan ukuran kamera standar. Ukuran layar dengan desain yang kecil menjadi pilihan responden sebab *smartphone* dapat dengan mudah dioperasikan menggunakan satu tangan dan juga *portable* (ringan dan dapat dibawa). Nilai kepentingan dari atribut kapasitas penyimpanan (*memory*) memiliki presentase sebesar 18,7 %. Hal ini menunjukkan bahwa pertimbangan responden sebelum memilih *smartphone Android* yang menjadi pilihan utama yaitu atribut kapasitas penyimpanan (*memory*).

**Kata kunci :** *Smartphone Android*, Analisis Konjoin, Nilai *Utility*, Nilai *importance*

### **I. PENDAHULUAN**

#### **1.1. Latar Belakang**

Ponsel adalah alat komunikasi yang terkoneksi jaringan *wireless* melalui gelombang radio yang menyediakan komunikasi suara, *Short Message Service* (SMS) dan *Multimedia Message Service* (MMS). Seiring perkembangan ponsel di era globalisasi seperti sekarang ini ponsel telah didukung dengan adanya layanan internet, seperti *browsing* dan *email*. Generasi ponsel dimulai semenjak adanya PDA (*Personal Digital Assistant*) yang memiliki kemampuan seperti membaca dan mengirim *email*, menjalankan berbagai macam aplikasi. Setelah generasi ponsel kemudian hadirlah generasi *smartphone* atau ponsel pintar yang memiliki kemampuan yang lebih canggih. Umumnya suatu ponsel dikatakan sebagai *smartphone* jika dapat berjalan pada perangkat lunak sistem

operasi yang lengkap. *Smartphone* merupakan ponsel sederhana dengan fitur canggih dan kemampuan mengirim dan menerima *email*, menjelajah internet, dan membaca *ebook*. *Smartphone* adalah miniatur komputer dengan kemampuan ponsel. *Smartphone* memiliki jenis yang berdasarkan sistem operasi seperti *windows phone*, *Symbian*, *Blackberry*, dan *Android* (Mayasari, 2012).

Perusahaan *smartphone* sudah banyak yang menggunakan sistem operasi yang berbasis *Android*. *Android* merupakan sistem perangkat *mobile* yang berbasis *linux* yang mencakup sistem operasi, *middleware* dan aplikasi. *Android* menyediakan *platform* terbuka bagi para pengembang untuk menciptakan aplikasi terbaru atau bersifat terbuka (*open source*) (Murtiwiayati dan Lauren, 2013).

Produk *smartphone Android* memiliki atribut yang menarik sehingga menjadi hal utama yang dilihat oleh konsumen sebelum memilih *Android*. Terutama pada kelengkapan sistem yang memberikan kenyamanan pada pengguna *smartphone* tersebut. Menurut Tjiptono dalam Atminingsih (2011), atribut produk adalah unsur-unsur produk yang dipandang penting oleh konsumen dan dijadikan dasar pengambilan keputusan. Atribut pada produk *smartphone Android* hanya dikaji secara spesifikasinya yaitu harga, kapasitas *memory (GB)*, Ukuran Kamera (MP), performa baterai (mAh), ukuran layar (inch) dan jaringan.

Menurut penelitian sebelumnya oleh Susanti (2008), metode penelitiannya menggunakan *cluster random sampling*, dimana sampel diambil dari Mahasiswa Matematika FMIPA Universitas Bengkulu yang memiliki *flash disk*. Preferensi ditentukan menggunakan analisis konjoin. Menggunakan metode *choice-based conjoint* atribut yang diamati adalah faktor merek, harga, kapasitas penyimpanan, fitur tambahan dan ukuran. Nilai atribut dari setiap level dapat diselesaikan menggunakan bantuan *software SPSS 16.0*.

Menurut Diana (2011), Atribut–atribut tersebut seringkali diamati baik secara keseluruhan maupun satu–persatu konsumen dalam melakukan penilaian terhadap suatu produk. Salah satu alat statistika yang dapat digunakan untuk menjawab permasalahan ini adalah statistik multivariat. Preferensi konsumen terhadap pemilihan atribut hal utama yang harus dilakukan adalah menentukan teknik analisis yang sesuai pada penelitian ini. Analisis yang digunakan adalah analisis konjoin (*conjoint analysis*) merupakan suatu metode

analisis dalam analisis multivariat, analisis konjoin mulai dikembangkan sejak tahun 1970. Analisis ini digunakan teruntuk membantu mendapatkan kombinasi dari atribut-atribut suatu produk atau jasa yang paling disukai konsumen. Dalam proses analisis konjoin akan memberikan ukuran kuantitatif terhadap tingkat kegunaan (*utility*) dan kepentingan relatif (*relatif importance*) suatu atribut dibandingkan dengan atribut lain.

Metode analisis yang terdapat pada analisis konjoin adalah metode *choice-based conjoint*. Karakteristik utama yang membedakan *choice-based conjoint* dari metode konjoin lainnya adalah responden menyatakan preferensi dengan melakukan pemilihan dari set konsep produk atau merancang kombinasi taraf dari atribut. Merancang kombinasi taraf atribut yang bisa disebut stimuli. Metode *choice-based conjoint* dapat dilakukan dengan pendekatan kombinasi lengkap (*full profile*) (Fauzy, 2012).

Berdasarkan latar belakang yang telah diuraikan diatas maka penulis tertarik menganalisis permasalahan tersebut menggunakan analisis konjoin untuk mengetahui preferensi mahasiswa Matematika FMIPA tahun 2011–2014 yang menggunakan *smartphone Android* yang ditinjau dari spesifikasi atribut yang digunakan yaitu harga, kapasitas *memory (RAM)*, Ukuran Kamera (MP), performa baterai (mAh), ukuran layar (Inch) dan jaringan. Sehingga berdasarkan uraian diatas maka penulis tertarik untuk mengetahui preferensi mahasiswa dalam memilih produk *smartphone Android* menggunakan analisis konjoin.

## II. LANDASAN TEORI

### 2.1. Analisis Konjoin

Analisis konjoin (*conjoint analysis*) merupakan suatu metode analisis dalam analisis multivariat yang biasa diterapkan pada riset pemasaran untuk menghasilkan suatu produk baru yang lebih disukai oleh konsumen. Analisis konjoin bertujuan untuk memberikan penilaian terhadap beberapa kombinasi taraf atribut yang ditawarkan. Penilaian yang diberikan oleh konsumen dapat diukur melalui peringkat (*rank*) atau skore (skala likert) (Mattjik dan Sumertajaya, 2011).

Menurut Hair *et al.* (2007), proses dari analisis konjoin akan memberikan ukuran kuantitatif terhadap tingkat kegunaan (*utility*) dan kepentingan relatif (*relatif importance*) suatu atribut produk. Analisis konjoin erat kaitannya dengan kombinasi dari taraf atribut. Kombinasi dari taraf atribut pada analisis konjoin dapat disebut dengan profil atau stimuli.

Hal utama didalam analisis konjoin, peneliti harus mengidentifikasi atribut dengan menentukan level atau taraf dari masing-masing atribut terlebih dahulu. Pembentukan suatu kombinasi taraf atribut (stimuli) yang biasa digunakan yaitu kombinasi lengkap (*full-profile*). Kombinasi lengkap (*full profile*) dapat mengevaluasi semua kombinasi atribut. Mengevaluasi semua kemungkinan kombinasi yang ada ternyata justru dapat menyulitkan responden, untuk itu dilakukan pengurangan terhadap kombinasi taraf. Salah satu pengurangan terhadap kombinasi taraf dapat dilakukan *orthogonal arrays*. *orthogonal arrays* merupakan bagian dari rancangan faktorial fraksional (*fractional factorial design*) pada perancangan ini akan diperoleh suatu kombinasi atribut yang hanya mengukur efek utamanya

saja, sementara interaksi antara atribut tidak terukur akan diabaikan, sehingga jumlah stimuli yang terbentuk akan jauh berkurang. Metode pada analisis konjoin yang menggunakan konsep stimuli *full-profile* adalah metode *choice-based conjoint* (Supranto, 2004).

#### 2.1.1. Metode Choice-Based Conjoint

Mattjik dan Sumertajaya (2011), metode yang digunakan untuk merancang kombinasi taraf dari atribut pada penelitian ini adalah dengan menggunakan *choice-based conjoint*. Karakteristik utama yang membedakan *choice-based conjoint* dengan metodologi lainnya adalah responden menyatakan preferensi dengan melakukan pemilihan dari set konsep produk. Keunggulan utama *choice-based conjoint* dibandingkan metode lain adalah prosedur pengumpulan datanya langsung mencerminkan perilaku pasar. Metode *choice-based conjoint*, tidak dapat digunakan dalam penelitian dalam jumlah yang banyak. Jumlah maksimum atribut yang dapat digunakan dalam konsep *full-profile* pada analisis konjoin adalah 6–10 atribut. Batas jumlah atribut yang efektif untuk metode *choice-based conjoint* adalah sedikit dari pada jumlah tersebut, karena responden harus memproses beberapa konsep *full-profile* secara bersamaan. Dalam *choice-based conjoint*, data dianalisis secara keseluruhan (*aggregate*). Sehingga tidak dapat menduga parameter atau nilai utilitas atribut pada tingkat individu. Pendekatan metode perancangan percobaan untuk *choice-based conjoint*, responden dihadapkan pada alternatif pilihan dari kombinasi atribut yang telah dirancang sebelumnya. Pada *choice-based Conjoint* merancang kombinasi atribut dapat menggunakan *full-profile*.

Menurut Bilshken dalam Anisyah (2014), metode *full-profile* merupakan metode yang paling realistis dalam menampilkan masing-masing taraf pada suatu profil dengan jelas. Semakin banyak jumlah atribut dan taraf, maka akan semakin banyak pula stimuli yang terbentuk.

### 2.1.2. Model Analisis Konjoin

Menurut Santoso (2004), Secara umum model dasar analisis konjoin untuk setiap faktor dan level ditulis dalam bentuk :

$$\mu(X) = \sum_{i=1}^m \sum_{j=1}^{k_i} a_{ij} X_{ij}$$

setiap atribut untuk setiap ( $i = 1, 2, 3, \dots, p$ )

$W_i$  = bobot kepentingan relatif untuk setiap atribut.

Setelah didapatkan nilai-nilai  $a_{ij}$  atau *utility*, maka kisaran *parth-worth*  $I_i$ , dan timbangan kepentingan relatif  $W_i$  akan diperoleh. Kisaran *parth-worth* dan timbangan kepentingan relatif ini memberikan dasar untuk menginterpretasikan hasil. Angka  $W_i$  yang terbesar menunjukkan preferensi terbesar terhadap level-level pada atribut tertentu.

## III. METODE PENELITIAN

### 3.1. Objek Penelitian

Pada penelitian ini yang menjadi objek penelitian adalah mahasiswa Matematika FMIPA Universitas Bengkulu Angkatan 2011-2014. Teknik yang digunakan pada penelitian ini yaitu teknik *nonprobability sampling*, dimana setiap unsur dalam populasi tidak memiliki kesempatan atau peluang yang sama untuk dipilih sebagai sampel. Cara penarikan sampel menggunakan *purposive sampling*, dimana responden memiliki karakteristik tertentu yang

akan dijadikan sampel, karakteristik yang dimiliki responden dalam penelitian ini adalah responden yang menggunakan *smartphone Android*. Metode yang digunakan untuk pengambilan sampel dari setiap kelompok (angkatan) menggunakan *cluster random sampling*.

### 3.2. Jenis Penelitian

Jenis penelitian yang digunakan adalah penelitian terapan yang menggunakan suatu metode dengan menerapkannya pada suatu kasus dan hanya berusaha memberikan gambaran atau mendeskripsikan objek atau permasalahan. Alat yang digunakan adalah kuesioner tertutup yang ditujukan kepada Mahasiswa Matematika FMIPA Universitas Bengkulu Angkatan 2011-2014.

### 3.3. Metode Pengumpulan Data

Metode pengumpulan data yang digunakan adalah dengan cara mendatangi secara langsung responden yang menjadi sampel dalam penelitian ini.

### 3.4. Populasi dan Sampel

Populasi dalam penelitian ini adalah Mahasiswa Matematika FMIPA Universitas Bengkulu Angkatan 2011-2014 yang masih aktif dan menggunakan produk *smartphone Android*. Populasi Mahasiswa Jurusan Matematika FMIPA Unib yang terdiri dari populasi disetiap Angkatan 2011 sebanyak 23 orang, 2012 sebanyak 37 orang, 2013 sebanyak 42 orang dan 2014 71 orang sesuai dengan data administrasi Fakultas Matematika dan Ilmu Pengetahuan Alam. Dari populasi tersebut, diambil sampel Mahasiswa Matematika FMIPA Unib Angkatan 2011-2014 yang menggunakan produk *smartphone Android*. Jumlah sampel ditentukan dengan rumus Slovin (Umar, 2008).

Menurut data administrasi Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Bengkulu, jumlah mahasiswa FMIPA Matematika Angkatan (2011-2014) yang masih aktif adalah 175 mahasiswa. Dengan menggunakan rumus *Slovin* ( $\epsilon = 5\%$ ) maka diperoleh jumlah sampel untuk responden adalah

$$n = \frac{175}{1 + 175 (0.05)^2} = 121,74 \approx 121$$

Populasi dari mahasiswa Matematika FMIPA Unib dari seluruh jumlah mahasiswa yang aktif, dapat dilihat dari kepemilikan mahasiswa matematika terhadap *smartphone Android*. Kemudian dari populasi tersebut dikelompokkan (*cluster*) berdasarkan setiap angkatan tahun 2011-2014. Elemen dari suatu anggota sampel dalam suatu angkatan memiliki sifat heterogen atau memiliki ciri yang berbeda pada kepemilikan *smartphone Android*. Teknik sampling yang digunakan pada penelitian ini adalah *cluster random sampling*. Menurut Cochran (1977), *cluster random sampling* adalah sampling terdiri dari kelompok atau cluster unit yang lebih kecil, yang biasa disebut elemen atau subunit populasi, kemudian sampel dipilih secara acak (*random*). Sampel memiliki sifat heterogen atau memiliki ciri unit sampel yang berbeda.

### 3.5. Sumber Data

Data primer dalam penelitian ini adalah data yang diperoleh secara langsung dari jawaban responden yang diteliti yaitu berupa jawaban responden yang diperoleh lewat pengisian kuesioner.

### 3.6. Rancangan Penelitian

Sebelum menyusun dan penulisan penelitian ini, penulis menyiapkan sejumlah Rancangan atau prosedur untuk menyusun skripsi penelitian ini.

Adapun rancangan tersebut terbagi menjadi beberapa tahap yaitu, yaitu :

Tahap identifikasi masalah Pada tahap ini penulis mencari inti permasalahan yang akan dibahas, yaitu Preferensi Mahasiswa dalam memilih produk *smartphone Android*.

- a) Tahap pengumpulan mengumpulkan materi-materi baik dari website, koran maupun buku-buku.
- b) Penyusunan kuisisioner
- c) Mencari alternatif pemecahan masalah Menganalisis secara Analisis Konjoin.
- d) Merumuskan masalah dengan teori, atribut yang dipilih harus sanga penting di dalam mempengaruhi preferensi dan pilihan responden.
- e) Membentuk stimuli/profil Pada penelitian ini, peneliti membentuk stimuli dengan desain profil penuh (*full-profile*) yang kemudian direduksi dengan cara *Orthogonal Array* melalui *software SPSS 16.0*.
- f) Menentukan bentuk data input Pada prosedur *full-profile* ini, para responden diminta untuk memberikan peringkat (*rank*) terhadap stimuli/kartu profil yang telah dibuat sebelumnya.
- g) Perhitungan tingkat kepentingan atribut Perhitungan tingkat kepentingan atribut dilakukan dengan menggunakan *software R 2.14.1*, pada saat melakukan analisis konjoin. Hasil pengukuran ini menunjukkan hubungan antara data dengan fakta yang ada dan juga keakuratan dalam memprediksi persepsi responden.

#### h) Interpretasi Hasil

Hasil interpretasi ini ditampilkan berupa plot-plot secara keseluruhan (*agregat*) yang kemudian dapat dideskripsikan supaya dapat dipahami oleh pembaca.

### 3.7. Variabel Penelitian

Variabel yang diamati adalah variable yang berupa:

1. Variabel yang terkait dalam perbedaan individu yaitu jenis kelamin, umur angkatan dan semester.
2. Variabel persepsi terhadap berbagai atribut pada *smartphone Android* yaitu
  1. Harga
  2. Kapasitas penyimpanan (GB)
  3. Ukuran kamera (MP)
  4. Performa baterai (mAh)
  5. Jaringan
  6. Ukuran layar (Inch)

## IV. HASIL DAN PEMBAHASAN

Pada bagian ini memuat hasil dan pembahasan analisis konjoin untuk mengetahui preferensi mahasiswa matematika FMIPA Unib angkatan 2011-2014 dalam memilih *smartphone Android*. Data yang digunakan adalah data yang diperoleh dari hasil kuisisioner mahasiswa Matematika FMIPA yang menggunakan *smartphone Android*

Setelah melakukan penelitian awal banyaknya sampel dari setiap Angkatan 2011-2014 yang menggunakan *smartphone Android* sesuai dengan hasil Slovin yaitu 121 responden. Sampel dibagi dari setiap angkatan 2011 sebanyak 17 orang, 2012 sebanyak 20 orang, 2013 sebanyak 30 orang dan 2014 sebanyak 54 orang.

## 4.1. Tahapan Analisis Konjoin

Analisis konjoin untuk mengetahui preferensi mahasiswa dalam memilih *smartphone Android* hal utama yang dilakukan adalah merancang kartu stimuli (kombinasi atribut produk). Kartu stimuli ini digunakan untuk mengetahui preferensi (penilaian) konsumen dalam memilih produk *smartphone Android* yang diinginkan oleh mahasiswa. Rancangan pada kartu stimuli ini akan digunakan sebagai kuisisioner penelitian. Adapun tahapan-tahapan analisis konjoin dengan menggunakan Program SPSS Versi 16.0 dan *software R 2.14.1*.

### 4.1.1. Merancang Kartu Stimuli

Tahapan yang dilakukan dalam analisis konjoin adalah membuat kombinasi dari taraf-taraf atribut produk atau merancang kartu stimuli. Hasil yang terbentuk dari suatu kombinasi atribut dapat dilihat pada Gambar. Banyaknya kombinasi yang terbentuk adalah 20 kombinasi.

No	HARGA	MEMORI	KAMERA	JARINGAN	BATERAI	LAYAR	STATUS	COMB.
1	+2.juta < 8 GB < 2000Mah < 4.	H3GA	5 MP < 5 MP4 inch < 5 inch	Design	1			
2	+2.juta < 8 GB < 4000Mah	H3GA	< 5 MP4 inch < 5 inch	Design	2			
3	1,5.juta < 2.juta < 8 GB < 10GB < 2000Mah	H3GA	5 MP < 8 MP < 5 inch	Design	3			
4	+2.juta < 32 GB < 2000Mah	4G LTE	5 MP < 8 MP < 4 inch	Design	4			
5	+1,5.juta < 8 GB < 10GB < 2000Mah	4G LTE	+8 MP4 inch < 5 inch	Design	5			
6	+2.juta < 8 GB < 10GB < 4000Mah	3G	+8 MP < 4 inch	Design	6			
7	+1,5.juta < 32 GB < 2000Mah < 4.	H3GA	+8 MP < 4 inch	Design	7			
8	1,5.juta < 2.juta < 32 GB < 2000Mah	3G	< 5 MP4 inch < 5 inch	Design	8			
9	+1,5.juta < 8 GB < 10GB < 4000Mah	4G LTE	< 5 MP < 5 inch	Design	9			
10	1,5.juta < 2.juta < 8 GB < 4000Mah	4G LTE	5 MP < 8 MP < 4 inch	Design	10			
11	+2.juta < 32 GB < 2000Mah < 4.	4G LTE	< 5 MP < 5 inch	Design	11			
12	+1,5.juta < 32 GB < 4000Mah	H3GA	5 MP < 8 MP4 inch < 5 inch	Design	12			
13	1,5.juta < 2.juta < 8 GB < 2000Mah < 4.	4G LTE	+8 MP4 inch < 5 inch	Design	13			
14	+1,5.juta < 8 GB < 2000Mah < 4.	3G	5 MP < 8 MP < 5 inch	Design	14			
15	1,5.juta < 2.juta < 8 GB < 10GB < 2000Mah < 4.	H3GA	< 5 MP < 4 inch	Design	15			
16	+1,5.juta < 8 GB < 2000Mah	3G	< 5 MP < 4 inch	Design	16			
17	1,5.juta < 2.juta < 32 GB < 4000Mah	3G	+8 MP < 5 inch	Design	17			
18	+2.juta < 8 GB < 2000Mah	H3GA	+8 MP < 5 inch	Design	18			
19	+2.juta < 8 GB < 2000Mah	H3GA	+8 MP < 4 inch	Hybrid	19			
20	1,5.juta < 2.juta < 8 GB < 4000Mah	3G	< 5 MP < 4 inch	Hybrid	20			

### 4.1.2. Analisis Konjoin

Analisis konjoin untuk menentukan nilai kepentingan pada suatu atribut dapat menggunakan *software R 2.14.1*. Setelah melakukan beberapa langkah yang diperlukan pada analisis ini didapatkan nilai *part-worth* atau nilai *utility* atau kegunaan dari setiap taraf atribut yang disukai oleh responden.

#### 4.2. Menentukan Kepentingan Atribut Menggunakan Software R 2.14.1

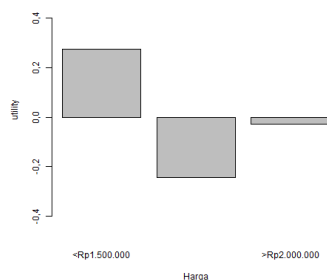
Dari 121 data responden hasil penyebaran kuisioner kepada mahasiswa Matematika FMIPA yang menggunakan *smartphone Android* diolah dengan menggunakan *software R 2.14.1*. Hasil analisis konjoin secara *aggregate (subfile Summary)* didapatkan nilai *utility* (kegunaan), nilai *importance* (kepentingan) dan gambar untuk mengetahui preferensi mahasiswa Matematika FMIPA dalam memilih *smartphone Android*.

##### 4.2.1. Nilai Utility (Kegunaan)

Nilai *utility* digunakan untuk menyatakan unilai kegunaan masing-masing level dalam faktor. Nilai ini digambarkan dalam bentuk grafik yang mewakili penilaian masing-masing responden. Apabila dalam grafik *utility* adalah positif, maka berarti responden tersebut menyukai taraf atribut tersebut, dan apabila negatif berarti responden tidak menyukai taraf atribut tersebut.

##### 1. Nilai Utility Atribut Harga

Nilai *utility* untuk atribut harga *smartphone* dapat dilihat pada Gamba Taraf atribut harga terbagi menjadi 3 taraf atribut yaitu taraf atribut dengan harga < Rp1.500.000, taraf atribut dengan harga Rp1.500.000 – Rp2.000.000 dan taraf atribut dengan harga > Rp2.000.000.

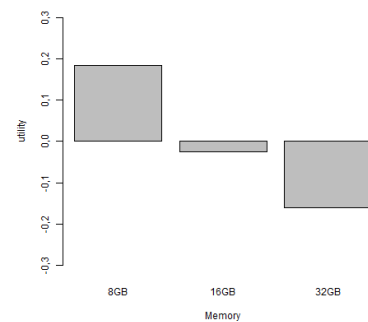


Taraf atribut dengan harga < Rp1.500.000 memiliki nilai *utility* sebesar

0,275, pada taraf atribut dengan harga Rp1.500.000 – Rp2.000.000 memiliki nilai *utility* –sebesar 0,2456 dan taraf atribut dengan harga > Rp2.000.000 memiliki nilai *utility* sebesar –0,0294.

##### 2. Nilai Utility Atribut Kapasitas Penyimpanan (memory)

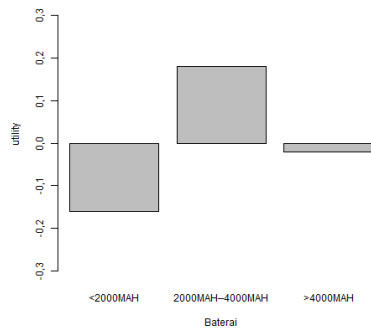
Kapasitas penyimpanan (*memory*) terbagi atas beberapa taraf atribut yaitu taraf atribut dengan kapasitas penyimpanan *memory* sebesar 8GB , 16GB dan 32GB.



Nilai *utility* yang memiliki kapasitas penyimpanan (*memory*) 8GB memberikan nilai *utility* sebesar 0,1856, kapasitas penyimpanan 16GB memberikan nilai *utility* sebesar 0,0253 dan 32GB memberikan nilai *utility* sebesar –0,1603. Hal ini menunjukkan bahwa nilai *utility* terbesar menunjukkan bahwa kecenderungan responden lebih menyukai kapasitas penyimpanan lebih tinggi yaitu *smartphone Android* yang memiki kapasitas sebesar 8GB.

##### 3. Nilai Utility Atribut Baterai (MaH)

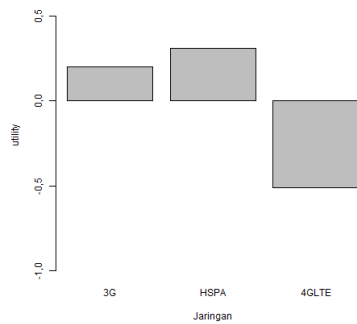
Atribut daya tahan baterai *smartphone* dalam satuan (mAh) dibagi dalam beberapa taraf atribut yaitu baterai kekuatan < 2000mAh, >4000mAh dan 2000mAh–4000 mAh.



Daya tahan baterai *smartphone* dalam satuan (mAh) dengan kekuatan <2000 mAh dan >4000mAh kurang diminati karena bernilai negatif  $-0,1602$  dan  $-0,0211$  sedangkan responden Lebih menyukai kekuatan 2000mAh-4000mAh menghasilkan nilai *utility* yang bernilai positif yaitu  $0,1813$ .

## 2. Nilai *Utility* Atribut Jaringan

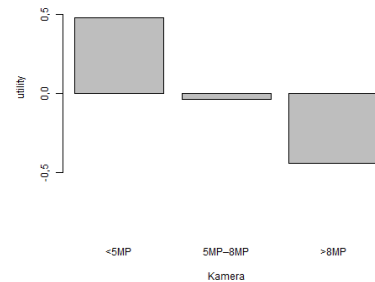
Atribut pada jaringan *smartphone Android* terbagi atas beberapa taraf yaitu 3G, HSPA dan 4G LTE.



Atribut pada jaringan *smartphone Android* 3G dan HSPA bernilai positif yaitu  $0,202$  dan  $0,3095$  sedangkan untuk jaringan 4GLTE nilai *utility* bernilai negatif yaitu  $0,5115$ . Hal ini menunjukkan bahwa responden lebih dominan menyukai jaringan 3G dan HSPA pada *smartphone* dibandingkan dengan jaringan 4G LTE.

## 3. Nilai *Utility* Atribut Kamera (MP)

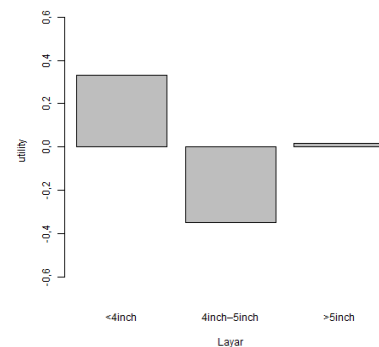
Ukuran kamera pada *smartphone Android* dalam satuan MP (Mega Pix) terbagi atas beberapa taraf atribut <5MP, 5MP-8MP dan >8MP.



Ukuran kamera pada taraf atribut <5MP memiliki nilai *utility* yang bernilai positif  $0,4789$  sedangkan 5MP-8MP dan >8MP memiliki nilai negatif  $-0,0377$  dan  $-0,4412$ .

## 4. Nilai *Utility* Atribut Layar (inch)

Atribut layar dalam satuan (inch) terbagi atas beberapa taraf atribut < 4inch >, 4inch-5inch dan >5inch.



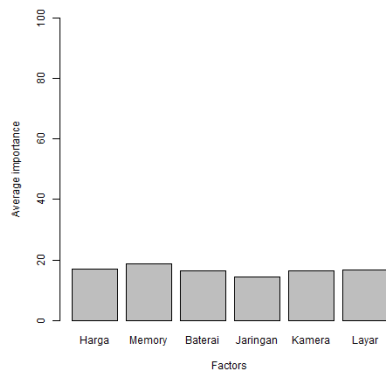
Atribut layar dalam satuan (inch) ukuran <4inch dan >5inch bernilai positif yaitu  $0,3316$  dan  $0,0174$ , untuk ukuran layar <4inch cenderung memberikan nilai terbesar. Sedangkan untuk ukuran layar 4inch-5inch memberikan nilai *utility* sebesar  $-0,349$ .

#### 4.3. Nilai *importance* (Kepentingan Atribut)

Nilai *importance* (kepentingan), yang mana nilai tersebut merupakan gabungan pendapat responden terhadap faktor yang dimaksud. Nilai *importance* digunakan untuk mengetahui faktor mana yang dianggap terpenting oleh responden dalam memilih produk *smartphone Android*.

Atribut	<i>Importance</i> %
<b>Harga</b>	17,05
<b>Memory</b>	18,7
<b>Baterai</b>	16,47
<b>Jaringan</b>	14,58
<b>Kamera</b>	16,55
<b>Layar</b>	16,65

Faktor kapasitas merupakan faktor pilihan yang pertama dalam memilih *smartphone Android*. Kemudian yang menjadi pertimbangan berikutnya adalah jaringan memiliki tingkat kesukaan yang rendah.



#### 4.4. Pembahasan Hasil Interpretasi

Hasil pembahasan dari analisis konjoin terhadap kecenderungan konsumen dalam memilih suatu atribut, dapat dilihat dari nilai kegunaan atau *utility* yang sudah dijelaskan pada Tabel 4.1 Sedangkan untuk melihat tingkat kepentingan atribut dapat dilihat pada Tabel 4.2.

## V. KESIMPULAN DAN SARAN

### 5.1. Kesimpulan

Beberapa kesimpulan yang dapat disampaikan yaitu preferensi dan tingkat kepentingan atribut yang menjadi pertimbangan responden dalam memilih produk *smartphone Android*.

1. Preferensi terhadap spesifikasi produk *smartphone Android*, dimana responden lebih memilih *smartphone Android* dengan harga lebih terjangkau, dengan ruang penyimpanan *memory* yang tidak terlalu besar, daya tahan baterai yang tahan lama sehingga dapat menunjang aktifitas pengguna *smartphone* dalam kebutuhan kesehariannya, kekuatan jaringan yang cepat, dan ukuran kamera standar. Ukuran layar yang kecil menjadi pilihan responden karena dapat dengan mudah dioperasikan menggunakan satu tangan dan *portable* (ringan dan dapat dibawa).
2. Berdasarkan hasil analisis konjoin bahwa atribut yang sangat mempengaruhi mahasiswa dalam memilih *smartphone Android* dapat dilihat dari nilai kepentingan (persentase) dari setiap atribut, dimana harga memiliki presentase sebesar 17,05 %, *memory* 18,7 %, baterai 16,47 %, jaringan 14,58 %, kamera 16,55 % dan layar 16,65 %. Hal ini menunjukkan bahwa pertimbangan responden sebelum memilih *smartphone Android* yang menjadi pilihan utama yaitu atribut kapasitas penyimpanan (*memory*), karena dapat dilihat dari jumlah persentase yang cukup besar yaitu 18,7%.

### 5.2. Saran

Saran yang dapat diberikan oleh penulis sebagai bahan penelitian lanjutan adalah analisis konjoin tidak hanya digunakan pada studi

pemasaran saja, akan tetapi dapat diterapkan pada berbagai bidang ilmu seperti psikologi, biologi, dan ilmu sosial lainnya. Metode yang dipilih oleh peneliti yaitu metode *choice based conjoint*, dimana responden memilih rancangan stimuli yang telah terbentuk sebelumnya, hal ini bertujuan untuk melihat tingkat preferensi dari setiap atribut yang cenderung dipilih oleh responden sebelum memilih atau membeli suatu produk *smartphone Android*, untuk penelitian selanjutnya metode *traditional conjoint* dan metode *adaptive conjoint analysis* (aca) dapat digunakan sebagai metode penelitian.

#### DAFTAR PUSTAKA

- Anisyah, N, 2014, Penerapan Analisis Konjoin untuk Keinginan Siswa Terhadap Pelayanan Kesehatan di Sma Harapan Medan Tahun 2013. *Skripsi*. Universitas Sumatra Utara. Medan.
- Atminingsih, 2012, Analisis Pengaruh Atribut Produk Sistem Operasi Ponsel Terhadap Sikap. *Skripsi*. Fakultas Manajemen. Universitas Gunadarma. Depok.
- Cochran, W, G, 1977, *Sampling Technniques*. Jhon Wiley dan Sons, Inc. New York.
- Diana, E. S, 2011, Pendekatan *Conjoint Analysis* untuk mengukur tingkat Preferensi Mahasiswa terhadap Layanan Perpustakaan UIN Sunan Kalijaga Yogyakarta. *Matematika FMIPA UNY Yogyakarta*.
- Fauzy, U, 2014, Aplikasi Analisis Konjoin dengan Model Regresi Logistik dalam Mengukur Preferensi Mahasiswa dalam memilih Ponsel. *Matematika*. Vol.3 No.1. Universitas Andalas. Padang.
- Joseph F. Hair JR, William C, Black, Barry J, Babin, Rolph E, Anderson 2007, *Multivariate analysis seventh Edition*, New Jersey : Prantice-Hall International.
- Mattjik, A. A dan Sumertajaya, M. I. 2011, Sidik Peubah Ganda dengan Menggunakan SAS. FMIPA IPB. Bogor.
- Mayasari, H, 2012, Analisis Perilaku Pembelian Ponsel Cerdas (Smartphone) antara Kebutuhan dan Gaya Hidup Konsumen di Kota Padang. *Ekonomi*. Vol.3No.1. Univeristas Taman Siswa. Padang.
- Murtiwiyati dan L. Lauren, 2013, Rancang Bangun Aplikasi Pembelajaran Budaya Indonesia untuk Anak Sekolah Dasar Berbasis Android. *Manajemen*. Vol.12 No. 2. Universitas Gunadarma. Depok.
- Santoso, S, 2003, *Buku Latihan SPSS Statistik Multivariat*. Elex Media Komputindo. Jakarta.
- Supranto, J, 2004, *Analisis Multivariat Arti dan interpretasi*. PT. Rineka Cipta. Jakarta.
- Susanti, N, 2008, Preferensi Mahasiswa dalam Pemilihan Flash Disk. *Skripsi*. FMIPA UNIB. Bengkulu.

**ANALISIS KORESPONDENSI TERHADAP VARIABEL FAKTOR YANG  
MEMPENGARUHI MINAT KONSUMEN DALAM BERBELANJA  
(STUDI KASUS : BENCOOLEN MALL)**

**Tiara Okti Miranti<sup>1</sup>, Sigit Nugroho<sup>2</sup>, dan Etis Sunandi<sup>2</sup>**

**1** Alumni Jurusan Matematika Fakultas MIPA Universitas Bengkulu

**2** Staf Pengajar Jurusan Matematika Fakultas MIPA Universitas Bengkulu

**ABSTRACT**

This study aims to determine the factor variable that affect consumer interest to shopping and. Correspondence Analysis (CA) is used in this study. The data are primary data in the form of questionnaires conducted on Bencoolen Mall. Research variable consisted of 10 indicator variables. The results showed that there are some factor variables need to be considered by the Manager of Bencoolen Mall to attract consumer interest in shopping, which are road access, product offering, pricing, salesman or saleswoman, facility, cleanliness and safety, service, diversity outlet and support. Every factor variables related closely, it indicates that each factor variable have the same effect for consumers.

Keyword: factors, *correspondence analysis*

**ABSTRAK**

Penelitian ini bertujuan untuk mengetahui variabel faktor yang mempengaruhi minat konsumen berbelanja dan keterkaitan setiap variabel faktor yang mempengaruhi minat konsumen dalam berbelanja. Analisis yang digunakan dalam penelitian adalah *Correspondence Analysis* (CA). Data yang digunakan adalah data primer berupa kuesioner yang dilakukan di Bencoolen Mall. Variabel penelitian terdiri dari 10 variabel indikator. Hasil penelitian menunjukkan bahwa ada beberapa variabel faktor yang harus dipertimbangkan oleh pihak pengelola Bencoolen Mall untuk menarik minat konsumen dalam berbelanja yaitu akses jalan, produk yang ditawarkan, harga, personil atau pramuniaga, fasilitas, kebersihan dan keamanan, pelayanan, keanekaragaman outlet dan sarana pendukung. Setiap variabel faktor mempunyai keterkaitan yang cukup erat, hal ini menunjukkan bahwa setiap variabel faktor mempunyai pengaruh yang sama bagi konsumen.

Kata kunci: faktor-faktor, analisis korespondensi

## **Pendahuluan**

Berkembangnya pasar modern saat ini memberikan alternatif bagi konsumen sebagai tempat berbelanja dibandingkan berbelanja di pasar tradisional.

Kota Bengkulu menjadi salah satu contoh kota yang menjadi tempat pembangunan pasar modern tersebut diantaranya *Bencoolen Mall*. Dalam penelitian ini dilakukan survei terhadap variabel faktor yang mempengaruhi minat konsumen dalam berbelanja.

Salah satu analisis multivariat yang digunakan dalam menentukan peta persepsi adalah analisis korespondensi (*correspondence analysis*).

Permasalahan dan tujuan dari penelitian ini adalah untuk menjelaskan penerapan analisis korespondensi terhadap variabel faktor dan untuk mengetahui keterkaitan antara setiap variabel faktor yang mempengaruhi minat konsumen dalam berbelanja.

## **Gambaran Umum Pasar Modern**

Pasar modern adalah pasar yang dikelola dengan manajemen yang modern, umumnya terdapat di kawasan perkotaan, sebagai penyedia barang dan jasa dengan mutu dan pelayanan yang baik kepada konsumen (umumnya anggota masyarakat kelas menengah ke atas). Pasar modern antara lain mall, supermarket, *departement store*, *shopping centre*, waralaba, toko mini swalayan, pasar serba ada, toko serba

ada dan sebagainya. Selanjutnya dari segi infrastruktur, pasar modern memiliki kelengkapan yang jauh lebih baik daripada pasar tradisional, seperti tersedianya penyejuk udara, tata letak yang memudahkan konsumen mencari barang, tersedianya eskalator dan lainnya (Natalia, 2010).

## **Faktor-Faktor Yang Mempengaruhi Konsumen**

Engel, Blackwell dan Miniard dalam Hilmi (2009), mengatakan ada beberapa faktor-faktor determinan dalam pemilihan toko. Faktor-faktor tersebut adalah sebagai berikut :

1. Lokasi
2. Kualitas keragaman
3. Harga
4. Iklan dan Promosi
5. Personil penjualan
6. Pelayanan yang ditawarkan
7. Atribut fisik dari toko
8. Sifat pelanggan toko
9. Suasana toko
10. Pelayanan dan kepuasan sesudah bertransaksi

## **Kuesioner**

Dalam Siregar (2013), Sukandarrumidi menyatakan bahwa kuesioner adalah alat pengumpulan data supaya diperolehnya informasi dengan cara memberikan daftar pernyataan tertulis kepada responden untuk diisi.

## Uji Validitas dan Uji Reliabilitas

Suatu tes atau instrumen pengukur dapat dikatakan mempunyai validitas yang tinggi apabila alat tersebut menjalankan fungsi ukurnya, atau memberikan hasil ukur yang sesuai dengan maksud dilakukannya pengukuran tersebut. Sedangkan Uji reliabilitas digunakan untuk mengetahui konsistensi alat ukur, apakah alat ukur yang digunakan dapat diandalkan dan tetap konsisten jika pengukuran tersebut diulang.

## Analisis Korespondensi

Analisis Korespondensi (*Correspondence Analysis*) merupakan bagian dari analisis multivariat yang mempelajari hubungan antara dua variabel atau lebih variabel dengan memperagakan baris dan kolom secara bersama dari tabel kontingensi dua arah dalam ruang vektor berdimensi rendah (Greenacre, 2007). Hasil dari analisis korespondensi biasanya menunjukkan dimensi terbaik untuk mempresentasikan data, yang menjadi koordinat titik dan suatu ukuran jumlah informasi yang ada dalam setiap dimensi yang disebut *inersia* (Johnson, 2002).

Jika  $X$  dan  $Y$  adalah dua peubah yang masing-masing mempunyai sebanyak  $a$  dan  $b$  kategori, maka dapat dibentuk suatu matriks data pengamatan  $P$  dengan ukuran  $a \times b$ . Dengan  $p_{ij} \geq 0$  menyatakan frekuensi dari sel ke  $(i, j)$ .

Berikut contoh matriks data  $P$  (Nugroho, 2008) :

$$P = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1b} \\ p_{21} & p_{22} & \dots & p_{2b} \\ \dots & \dots & \ddots & \vdots \\ p_{a1} & p_{a2} & \dots & p_{ab} \end{bmatrix} \quad (2.3)$$

Dari matriks data  $P$  di atas dapat dibentuk tabel kontingensi dua arah sebagai berikut :

**Tabel 2.1. Tabel kontingensi dua arah**

	$Y_1$	$\dots$	$Y_j$	$\dots$	$Y_b$	Total
$X_1$	$p_{11}$	$\dots$	$p_{1j}$	$\dots$	$p_{1b}$	$p_{1.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$X_i$	$p_{i1}$	$\dots$	$p_{ij}$	$\dots$	$p_{ib}$	$p_{i.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$X_a$	$p_{a1}$	$\dots$	$p_{aj}$	$\dots$	$p_{ab}$	$p_{a.}$
Total	$p_{.1}$	$\dots$	$p_{.j}$	$\dots$	$p_{.b}$	$p_{..}$

Keterangan :

$$p_{i.} = \sum_{j=1}^b p_{ij}, \text{ peluang marjinal } X$$

$$p_{.j} = \sum_{i=1}^a p_{ij}, \text{ peluang marjinal } Y$$

$$p_{..} = \sum_i \sum_j p_{ij}, \text{ jumlah total frekuensi dari matriks } P$$

$p_{ij}$ , frekuensi pengamatan ke  $i$  baris pada  $j$  kolom

Uji yang sesuai untuk mengetahui ada tidaknya hubungan antara dua variabel kategori yang berupa tabel kontingensi, adalah *Pearson Chi-Square test*, statistik ujinya adalah (Johnson, 2002):

$$\chi^2 = \sum_{i=1}^a \sum_{j=1}^b \frac{(p_{ij} - m_{ij})^2}{m_{ij}} \quad (2.4)$$

Keterangan:

$p_{ij}$  = Jumlah pengamatan pada baris ke- $i$  dan kolom ke- $j$

$m_{ij}$  = Frekuensi harapan

$a$  = Banyaknya baris

$b$  = Banyaknya kolom

### Metode Penelitian

Penelitian ini dilakukan dalam bentuk studi kasus dengan panduan atau referensi dari berbagai buku teks dan jurnal. Analisis data menggunakan software R.3.2.5 dan SPSS sebagai input data. Adapun tahapan penelitian sebagai berikut:

1. Pengumpulan data
2. Penyusunan dan penyebaran kuesioner
3. Uji validitas dan uji reliabilitas
4. Menyusun tabel kontingensi
5. Uji *Chi-Square* ( $\chi^2$ )

Uji ini berguna untuk mengetahui hubungan antara penilaian responden dengan setiap variabel faktor.

6. Analisis Korespondensi  
Tahap-tahap dalam analisis korespondensi sebagai berikut :

- a. Dari tabel kontingensi data asal disusun kedalam bentuk matrik dan dilakukan penguraian nilai singular untuk mengetahui nilai variabilitas data asli yang dijelaskan oleh setiap dimensi yang dihasilkan.

- b. Melakukan analisis korespondensi pada tabel kontingensi dengan bantuan *software R.3.2.5*.

- c. Mengamati nilai koordinat dan visualisasi plot profil vektor baris dan kolom dalam setiap titik yang terdekat pada masing-masing segmen untuk mendeskripsikan minat berbelanja.

### Hasil dan Pembahasan

Analisis data ditempuh dengan melakukan beberapa tahapan, yakni, uji data dan analisis korespondensi dengan R.3.2.5.. Uji data dilakukan untuk mengetahui validitas dan reabilitas data. Uji validitas dan reabilitas data dilakukan dengan menggunakan uji *cronbach alpha*. Semua data yang lolos uji validitas dan reabilitas dapat digunakan untuk analisis selanjutnya. Analisis selanjutnya adalah analisis data dengan menggunakan *software R.3.2.5*. Tujuan analisis untuk menjelaskan penerapan analisis korespondensi terhadap variabel faktor yang mempengaruhi minat konsumen dalam berbelanja dan mengetahui keterkaitan antara setiap variabel faktor yang mempengaruhi minat konsumen dalam berbelanja.

Tahapan awal dalam menganalisis data adalah dengan merumuskan hipotesis untuk uji *Chi-Square* ( $\chi^2$ ) yaitu :

$H_0$ : tidak ada hubungan antara penilaian responden dengan variabel faktor

$H_1$ : ada hubungan antara penilaian responden dengan variabel faktor

Dari Tabel 4.7 dapat digunakan untuk mencari nilai *chi-square* dengan menggunakan persamaan 2.4, sehingga diperoleh sebagai berikut:

$$\chi^2 = \sum \frac{(D_{ij} - m_{ij})^2}{m_{ij}}$$

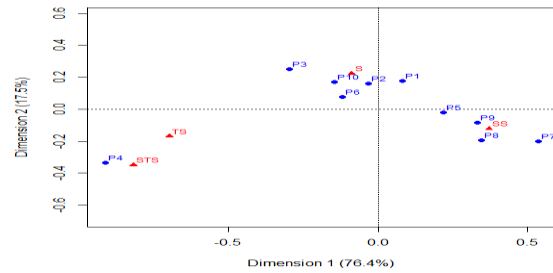
$$\chi^2 = 401,354$$

Nilai *chi-square* ( $\chi^2$ ) dengan  $\alpha = 0,05$  adalah 40,113 dan karena  $40,113 < 401,354$  maka  $H_0$  ditolak yang artinya ada hubungan signifikan antara penilaian responden dengan variabel faktor.

Selanjutnya dilakukan perhitungan analisis CA yang dilakukan dengan bantuan program R (Lampiran 7) sehingga diperoleh informasi bahwa terdapat tiga *eigenvalue* sebagai berikut.

1. Faktor pertama *eigenvalue* sebesar 0,153 mampu menerangkan keragaman data sebesar 76,43%
2. Faktor kedua *eigenvalue* sebesar 0,035 mampu menerangkan keragaman data sebesar 17,54% (total keragaman dua faktor pertama adalah 93,97%)
3. Faktor ketiga *eigenvalue* sebesar 0,012 mampu menerangkan keragaman data sebesar 6,03% (total keragaman tiga faktor pertama adalah 100%)

Dari hasil analisis CA yang dilakukan diperoleh *plotting* sebagai berikut.



**Gambar 4.1** *Plotting* variabel faktor terhadap penilaian responden

### Kesimpulan dan Saran

Berdasarkan hasil penelitian yang dilakukan, responden berpendapat “Sangat Setuju” bahwa variabel pramuniaga atau personil, kebersihan dan keamanan, pelayanan yang diberikan, dan keanekaragaman outlet sangat mempengaruhi minat konsumen dalam berbelanja. Sedangkan untuk pendapat “Setuju”, responden menunjukkan bahwa variabel akses jalan, produk yang ditawarkan, harga yang ditetapkan, fasilitas, dan sarana pendukung juga dapat mempengaruhi minat konsumen dalam berbelanja. Kemudian, responden berpendapat untuk “Tidak Setuju” dan “Sangat Tidak Setuju” bahwa variabel iklan dan promosi tidak mempengaruhi minat konsumen dalam berbelanja. Oleh karena itu, variabel-variabel yang dapat dipertahankan oleh pihak pengelola Bencoolen Mall untuk menarik minat konsumen dalam berbelanja yaitu pramuniaga atau personil, kebersihan dan keamanan, pelayanan yang diberikan, keanekaragaman outlet, akses jalan, produk yang ditawarkan, harga yang ditetapkan, fasilitas, dan sarana

pendukung. Dan setiap variabel faktor mempunyai keterkaitan yang cukup erat, sehingga setiap variabel faktor mempunyai pengaruh yang sama bagi penilaian konsumen.

## DAFTAR PUSTAKA

- Agresti, A. 1990. *Categorical Data Analysis*. John Wiley and Sons. New York.
- Azwar, S. 1997. *Reliabilitas dan Validitas*. Pustaka Pelajar. Jakarta.
- Greenacre, Micheal.J. 2007. *Correspondence Analysis In Practice, 2<sup>th</sup> Edition*. Universitat Pompeu Fabra Barcelona. Spain.
- Hilmi. 2009. *Analisis Faktor-Faktor yang Mempengaruhi Keputusan Konsumen Memilih Restoran Paparon Pizza di Kota Lhokseumawa*. Jurnal Ekonomi dan Bisnis Vol 9, No. 2 (2009).
- Johnson, Richard.A and Dean W.Wichern. 2002. *Applied Multivariate Statistical Analysis, 5<sup>th</sup> Edition*. Practice Hall Inc. New Jersey.
- Juliandi, A. 2007. *Teknik Pengujian Validitas Dan Reliabilitas*. Diakses 25 Februari 2016. [www.azuarjuliandi.com](http://www.azuarjuliandi.com)
- Khomariah, N. 2014. *Analisis Perbandingan Motivasi Konsumen Berbelanja di Pasar Tradisional Kemiri Muka Dengan Pasar Modern Carrefour ITC Depok*. Skripsi. Fakultas Ekonomi Universitas Gunadarma. Jakarta.
- Kotler, Phillip dan Kevin Lane Keller. 2009. *Manajemen Pemasaran Edisi 12*. PT.Indeks. Jakarta.
- Mulyadi, Raf. 2012. *Pengaruh yang Memotivasi Konsumen Berbelanja terhadap Keputusan Konsumen Berbelanja di Kota Jambi*. Jurnal Manajemen. Vol 1 No.1 Januari – Maret 2012, hal 66-68.
- Natalia, L. 2010. *Analisis Faktor Persepsi yang Mempengaruhi Minat Konsumen Untuk Berbelanja Pada Giant Hypermarket Bekasi*. Jurnal Skripsi, Fakultas Ekonomi Universitas Gunadarma. Jakarta. Diakses 10 Desember 2015. [http://www.gunadarma.ac.id/library/articles/graduate/economy/2009/Artikel\\_102\\_05725.pdf](http://www.gunadarma.ac.id/library/articles/graduate/economy/2009/Artikel_102_05725.pdf)
- Nugroho, S. 2008. *Statistika Multivariat Terapan*. UNIB Press. Bengkulu.
- Pangastuti, A. 2013. *Pemetaan Persepsi Merk Laptop di Kalangan Mahasiswa Menggunakan Analisis Korespondensi Berganda (Studi Kasus:Mahasiswa Universitas Diponegoro Semarang)*. Skripsi, Fakultas Sains dan Matematika Universitas Diponegoro. Semarang. Diakses 25 November 2015. [http://eprints.undip.ac.id/42344/1/Anissa\\_Pangastuti.pdf](http://eprints.undip.ac.id/42344/1/Anissa_Pangastuti.pdf)
- Rencher, C.Alvin. 2002. *Methods of Multivariate Analysis Second Edition*. A John Wiley and Sons. Inc.Publication. Canada.
- Rosalina, N.Erin. 2013. *Analisis Korespondensi Sederhana dan Berganda Pada Bencana Alam Klimatologis di Pulau Jawa*. Skripsi, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember. Jember.
- Singarimbun, M., Effendi, S. 1987. *Validitas dan reliabilitas Instrumen Penelitian Dalam Djamaludin Acok ed. Metode penelitian survei, 122-124*. LP3S. Jakarta.
- Siregar, H. 2013. *Analisis Persepsi Masyarakat Terhadap Terjadinya Kecelakaan Sepeda Motor di Kota Bengkulu*. Skripsi, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Bengkulu. Bengkulu.
- Wahyudi, R. 2009. *Uji validitas dan reliabilitas dengan pendekatan konsistensi internal kuesioner pembukaan program studi statistika FMIPA Universitas Bengkulu*. Skripsi, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Bengkulu. Bengkulu.
- Wishman, B.S.S. 2009. *Persepsi Konsumen Terhadap Prioritas Perbaikan Infrastruktur Pasar Perumnas Klender*. Tesis, Fakultas Ekonomi Universitas Indonesia, Jakarta.

**ANALISIS PENERIMA RASKIN DAN KARTU INDONESIA SEHAT (KIS)  
DENGAN MENGGUNAKAN REGRESI LOGISTIK BINER BIVARIAT  
(Studi Kasus: Rumah Tangga Miskin Penerima Raskin dan KIS Desa Talang  
Pauh Tahun 2016)**

**ANALYSIS OF RASKIN AND KARTU INDONESIA SEHAT (KIS) RECIPIENTS  
BY USING BIVARIATE BINARY LOGISTIC REGRESSION  
(Case Study: Poor Households Raskin and KIS recipients Talang Pauh  
Village in 2016)**

Bintari Wulan Sawitri<sup>1</sup>, Sigit Nugroho<sup>2</sup>, Dian Agustina<sup>3</sup>

Mahasiswa Jurusan Matematika FMIPA, Universitas Bengkulu<sup>1</sup>

Dosen Jurusan Matematika FMIPA, Universitas Bengkulu<sup>2,3</sup>

Email : [bbintarii@gmail.com](mailto:bbintarii@gmail.com)<sup>1</sup>

**ABSTRAK**

Pemerintah Indonesia selalu mengupayakan penanggulangan kemiskinan salah satunya dengan meluncurkan program Beras Miskin (Raskin) dan program Kartu Indonesia Sehat (KIS). Penelitian ini bertujuan untuk mendeskripsikan karakteristik rumah tangga miskin penerima Raskin dan KIS serta menentukan model rumah tangga miskin penerima Raskin dan KIS di Desa Talang Pauh. Analisis yang digunakan adalah regresi logistik biner bivariat dimana penerima Raskin dan KIS Desa Talang Pauh Tahun 2016 sebagai variabel respon dan 14 indikator kriteria rumah tangga miskin yang telah ditetapkan oleh BPS sebagai variabel prediktor dengan *Maximum Likelihood Estimation* (MLE) sebagai metode estimasi parameternya. Pada penelitian ini diperoleh 4 variabel prediktor yang dapat digunakan dalam analisis. Analisis menunjukkan bahwa variabel jenis lantai berpengaruh signifikan terhadap model regresi logistik biner bivariat dengan penerima Raskin dan KIS sebagai variabel respon. influence Raskin and KIS recipients in the bivariate binary logistic regression model.

**Kata kunci:** regresi logistik biner bivariat, *Maximum Likelihood Estimation*, Raskin, KIS

**ABSTRACT**

The Indonesian government is seeking to reduce poverty in Indonesia, one of which is launching Beras Miskin program (Raskin) and Kartu Indonesia Sehat program (KIS). This research aims to describe the characteristics of poor households Raskin and KIS recipients and determine a model of poor households Raskin and KIS recipients in Talang Pauh. The analysis used is bivariate binary logistic regression where Raskin and KIS recipients in Talang Pauh as the response variable and the 14 indicators of poor household criteria set by the BPS as predictor variable with Maximum Likelihood Estimation (MLE) as the parameter estimation method. In this research, four predictor

variables that can be used in the analysis. The result shows that floor type significantly influence Raskin and KIS recipients in the bivariate binary logistic regression model.

**Keywords:** bivariate binary logistic regression, Maximum Likelihood Estimation, Raskin, KIS

## 1. PENDAHULUAN

Kemiskinan merupakan salah satu permasalahan besar yang dihadapi oleh bangsa Indonesia. Pemerintah Indonesia selama ini selalu mengupayakan penanggulangan kemiskinan dengan meluncurkan program-program penanggulangan kemiskinan. Salah satu program pemerintah adalah program subsidi beras bagi masyarakat berpendapatan rendah atau yang biasa dikenal dengan Raskin. Program lainnya yaitu program yang melayani peserta Kartu Indonesia Sehat (KIS). Kedua program ini diperuntukkan bagi keluarga kurang mampu. BPS telah menetapkan kriteria rumah tangga miskin dimana ada 14 indikator yang digunakan untuk menentukan rumah tangga miskin. Penelitian ini dilakukan untuk melihat apakah suatu rumah tangga miskin terdaftar atau tidak sebagai penerima program Raskin dan Kartu Indonesia Sehat (KIS) sebagai variabel respon, sehingga ada dua kategori (biner) pada variabel responnya. Variabel prediktor yang digunakan berupa kriteria kemiskinan yang telah ditetapkan oleh BPS. Oleh karena itu, penelitian ini akan menggunakan metode pendekatan model logit biner bivariat dengan *Maximum Likelihood Estimation* (MLE) sebagai metode pendugaan parameternya.

## 2. METODE PENELITIAN

Data yang digunakan dalam penelitian ini adalah data sekunder berupa data rumah tangga miskin penerima Raskin dan KIS tahun 2016 Desa Talang Pauh dan data primer berupa data yang dikumpulkan dari kuesioner terhadap rumah tangga miskin penerima Raskin dan KIS Desa Talang Pauh. Teknik pengumpulan data yang digunakan adalah *Purposive Sampling*. Ukuran sampel yang diambil adalah rumah tangga miskin di Desa Talang Pauh yang menerima Raskin dan KIS tahun 2016 sebanyak 106 rumah tangga. Variabel yang digunakan dalam penelitian ini adalah dua variabel respon dan 14 variabel prediktor. Variabel respon yang digunakan beserta kategorinya disajikan dalam tabel berikut.

Tabel 2.1. Variabel respon yang digunakan

Variabel	Nama Variabel	Kategori
$Y_1$	Penerima Raskin	1 : Ya 0 : Tidak
$Y_2$	Penerima KIS	1 : Ya 0 : Tidak

Variabel prediktor yang digunakan beserta kategorinya disajikan dalam tabel berikut.

Tabel 2.2. Variabel prediktor yang digunakan

Variabel	Nama Variabel	Kategori
$X_1$	Luas bangunan tempat tinggal	1 : $< 8m^2$ 2 : $\geq 8m^2$
$X_2$	Jenis lantai tempat tinggal	1 : tanah 2 : bambu / kayu 3 : ubin
$X_3$	Jenis dinding tempat tinggal	1 : bambu 2 : rumbia / kayu 3 : tembok
$X_4$	Fasilitas buang air besar	1 : bersama 2 : milik sendiri
$X_5$	Sumber penerangan	1 : listrik 2 : bukan listrik
$X_6$	Sumber air minum	1 : sumur 2 : sungai 3 : air hujan
$X_7$	Bahan bakar memasak sehari-hari	1 : kayu bakar 2 : minyak tanah 3 : gas
$X_8$	Konsumsi daging/susu/ayam	1 : satu kali seminggu 2 : $>$ satu kali seminggu
$X_9$	Membeli pakaian baru dalam setahun	1 : satu pasang pakaian 2 : $>$ satu pasang
$X_{10}$	Makan dalam sehari	1 : dua kali 2 : $>$ dua kali
$X_{11}$	Tindakan pengobatan jika sakit	1 : obat warung 2 : puskesmas 3 : bidan desa
$X_{12}$	Penghasilan perkapita perbulan	1 : $<$ Rp. 600.000,- 2 : $\geq$ Rp. 600.000,-
$X_{13}$	Pendidikan kepala rumah tangga	1 : SD 2 : SMP 3 : SMA
$X_{14}$	Memiliki barang yang mudah dijual	1 : ada 2 : tidak ada

Sumber : Analisis dan Perhitungan Tingkat Kemiskinan oleh BPS Tahun 2008.

Tahapan analisis yang dilakukan agar tercapainya tujuan dari penelitian ini adalah sebagai berikut.

- a. Mendeskripsikan karakteristik rumah tangga miskin penerima program Raskin dan Kartu Indonesia Sehat (KIS) tahun 2016 di Desa Talang Pauh.
  - 1) Menganalisa variabel respon dan prediktor menggunakan analisis deskriptif.
  - 2) Menginterpretasikan hasil analisis deskriptif untuk memperoleh karakteristik rumah tangga miskin penerima Raskin dan KIS tahun 2016 di Desa Talang Pauh.
- b. Menentukan model rumah tangga miskin penerima program Raskin dan Kartu Indonesia Sehat (KIS) di Desa Talang Pauh menggunakan regresi logistik biner bivariat.
  - 1) Melakukan uji independensi antar variabel respon  $Y_1$  dan  $Y_2$  dengan menggunakan uji kai-kuadrat.
  - 2) Meregresikan semua variabel prediktor untuk memperoleh model regresi logistik biner bivariat awal kemudian dilakukan pengujian parameter secara parsial menggunakan uji *Wald*.
  - 3) Meregresikan variabel-variabel prediktor yang signifikan secara serentak menggunakan uji *likelihood ratio test* untuk memperoleh model regresi logistik biner bivariat yang baru.
  - 4) Menginterpretasikan model regresi logistik biner bivariat dan mengambil kesimpulan.

### 3. HASIL DAN PEMBAHASAN

#### 3.1. Karakteristik Rumah Tangga Miskin Desa Talang Pauh

Desa Talang Pauh dihuni oleh 484 Kepala Keluarga yang terbagi dalam 3 dusun. Dusun 1 dihuni sebanyak 164 Kepala Keluarga, sedangkan Dusun 2 dihuni sebanyak 283 Kepala Keluarga, dan sisanya 37 Kepala Keluarga berpenghuni di Dusun 3. Di Desa Talang Pauh terdapat keluarga atau rumah tangga yang masuk dalam kriteria penerima program pemerintah khususnya program Raskin dan KIS. Daftar penerima Raskin sebanyak 82 rumah tangga dan penerima KIS sebanyak 79 rumah tangga yang terbagi dalam Dusun 1, Dusun 2, dan Dusun 3. Persentase Penerima Raskin sebesar 25.5% yang artinya sebanyak 27 rumah tangga yang menerima Raskin saja dan sebesar 74.5 % yang artinya sebanyak 79 rumah tangga yang tidak menerima Raskin dari 106 rumah tangga penerima Raskin dan penerima KIS. Persentase Penerima KIS sebesar 22.6% yang artinya sebanyak 24 rumah tangga yang menerima KIS saja dan sebesar 77.4 % yang artinya sebanyak 82 rumah tangga yang tidak menerima KIS dari 106 rumah tangga penerima Raskin dan penerima KIS.

Di Desa Talang Pauh dari 106 rumah tangga miskin penerima Raskin dan KIS, tidak ada rumah tangga yang memiliki luas bangunan tempat tinggal kurang dari 8m<sup>2</sup>. Jenis lantai tempat tinggal dari tanah sebesar 68.9%, 0% untuk jenis lantai tempat tinggal dari bambu atau kayu, dan 31.1% untuk jenis lantai tempat tinggal dari ubin. Jenis dinding dari bambu atau rumbia sebesar 0%, 39.6% untuk jenis dinding dari kayu, dan 60.4% untuk jenis dinding dari tembok. Fasilitas buang air besar milik sendiri dan sumber penerangan menggunakan listrik. Sumber air minum berasal dari sumur dan bahan bakar memasak sehari-hari menggunakan gas. Rumah tangga miskin penerima Raskin dan KIS Desa Talang Pauh dalam mengkonsumsi daging atau susu atau ayam lebih dari satu kali dalam seminggu dan membeli lebih satu stel pakaian dalam setahun serta makan sebanyak lebih dari dua kali dalam sehari. Tindakan pengobatan jika sakit yaitu mengkonsumsi obat warung jika sakit sebesar 41.5%, untuk tindakan pengobatan di puskesmas sebesar 43.4%, untuk tindakan pengobatan bidan desa sebesar 15.1% dari 106 rumah tangga miskin penerima Raskin dan KIS Desa Talang Pauh. Penghasilan lebih dari Rp.600.000,- per bulan dan dari 106 rumah tangga miskin penerima Raskin dan KIS, kepala rumah tangga dengan pendidikan terakhir SD sebesar 39.6%, kepala rumah tangga dengan pendidikan terakhir SMP sebesar 47.2%, kepala rumah tangga dengan pendidikan terakhir SMA sebesar 13.2%, dengan memiliki tabungan atau barang yang mudah dijual. Persentase variabel-variabel prediktor tersebut memberikan informasi bahwa tidak semua rumah tangga miskin penerima program Raskin dan KIS memenuhi ke-empatbelas kriteria rumah tangga miskin yang ditetapkan oleh BPS. Hal ini berarti kurangnya ketepatan sasaran program Raskin dan KIS pada rumah tangga miskin Desa Talang Pauh.

### 3.2. Uji Independensi

Uji Independensi dilakukan untuk melihat keterkaitan antara variabel  $Y_1$  dan  $Y_2$ . Dengan menggunakan SPSS 16 diperoleh nilai  $\chi^2 = 10.60 > \chi^2_{(0.05;1)} = 3.84$  atau nilai  $sig = 0.001 < \alpha = 0.05$  artinya ada hubungan antara variabel respon  $Y_1$  dan  $Y_2$  sehingga dapat dilakukan analisis regresi logistik biner bivariat.

### 3.3. Uji Validitas dan Reliabilitas

Penelitian ini menggunakan kuisioner yang terdiri dari 14 butir pertanyaan dengan 40 responden sebagai pengujian awal. Dalam penelitian ini nilai r-tabel yang digunakan yaitu 0.312 dengan derajat bebas  $n = 40$  dan  $\alpha = 5\%$ . Uji validitas dan reliabilitas dilakukan menggunakan SPSS 16. Pertanyaan 2, 3, 11, dan 13 valid sedangkan yang lainnya tidak valid. Uji validitas lanjutan menunjukkan bahwa butir

pertanyaan 2, 3, 11, dan 13 sudah valid. Maka butir pertanyaan tersebut dapat digunakan dalam kuisisioner dan dapat digunakan sebagai variabel prediktor untuk dilakukan analisis. Uji reliabilitas juga dilakukan dengan SPSS 16 yang menunjukkan nilai *Cronbach's Alpha* = 0.864 > 0.6 artinya kuisisioner dengan 4 butir pertanyaan tersebut sudah reliabel dan dapat digunakan untuk analisis.

### 3.4. Analisis Regresi logistik Biner Bivariat

Analisis regresi logistik bivariat menggunakan software R 3.2.2 untuk memperoleh parameter-parameter dalam analisis regresi logistik biner bivariat.

Tabel 3.1. Pengujian regresi logistik biner bivariat secara individu

Variabel	Parameter	Koefisien	Standar error	Wald	p-value
Konstanta	$\beta_{01}$	3.0804	1.7719	1.738	0.0821
	$\beta_{02}$	1.3964	2.2086	0.632	0.5272
	$\gamma_0$	5.4987	100.3589	0.055	0.9563
Jenis lantai	$\beta_{11}$	0.1961	0.3639	0.539	0.5899
	$\beta_{12}$	-1.6021	0.5343	-2.999	0.0027
	$\gamma_1$	3.2588	108.0358	0.030	0.9759
Jenis dinding	$\beta_{21}$	-0.4147	1.2269	-0.338	0.7353
	$\beta_{22}$	1.2763	1.3843	0.922	0.3565
	$\gamma_2$	7.0650	28.3566	0.249	0.8032
Tindakan pengobatan	$\beta_{31}$	-0.4604	0.5354	-0.860	0.3898
	$\beta_{32}$	1.5314	0.9116	1.680	0.0929
	$\gamma_3$	-9.2807	191.4264	-0.048	0.9613
Pendidikan terakhir	$\beta_{41}$	-0.2269	1.1556	-0.196	0.8443
	$\beta_{42}$	-1.7887	1.2759	-1.402	0.1609
	$\gamma_4$	-14.6806	48.3782	-0.303	0.7615

Tabel 3.1 menunjukkan variabel jenis lantai tempat tinggal ( $X_2$ ) dengan nilai  $W = -2.999 > Z_{\alpha/2} = -1.6449$  atau dapat dilihat juga dari nilai  $p - value = 0.0027 < \alpha = 0.05$  yang artinya variabel jenis lantai tempat tinggal ( $X_2$ ) berpengaruh signifikan pada taraf  $\alpha = 5\%$ . Sedangkan variabel jenis dinding tempat tinggal ( $X_3$ ), tindakan pengobatan jika sakit ( $X_{11}$ ), dan pendidikan terakhir kepala rumah tangga ( $X_{13}$ ) tidak berpengaruh signifikan pada taraf  $\alpha = 5\%$ .

Pengujian regresi logistik biner bivariat secara serentak dilakukan dengan meregresikan seluruh variabel prediktor dengan variabel respon dan diperoleh nilai *likelihood ratio test*  $G = 107.1043 > \chi_{0.05;12}^2 = 21.0261$ . Hal ini menunjukkan bahwa variabel jenis lantai ( $X_2$ ) dapat digunakan dalam model akhir. Model regresi logistik biner bivariat yang dapat dibentuk dari Tabel 4.7 yaitu model logit:

$$g_1(x) = \ln \left[ \frac{\pi_1(x)}{1 - \pi_1(x)} \right] = 3.0804 + 0.1961 X_2$$

$$g_2(x) = \ln \left[ \frac{\pi_2(x)}{1 - \pi_2(x)} \right] = 1.3964 - 1.6021 X_2$$

Model peluang marginal untuk  $Y_1$  dan  $Y_2$  adalah :

$$\pi_1(x) = \frac{\exp(3.0804 + 0.1961 X_2)}{1 + \exp(3.0804 + 0.1961 X_2)}$$

$$\pi_2(x) = \frac{\exp(1.3964 - 1.6021 X_2)}{1 + \exp(1.3964 - 1.6021 X_2)}$$

Interpretasi model regresi biner bivariat dengan variabel respon Penerima raskin dan KIS serta variabel prediktor jenis lantai tempat tinggal ditampilkan pada Tabel 3.2.

Tabel 3.2. Interpretasi model regresi biner bivariat dengan odds rasio.

Kategori	Variabel	Odds Rasio	Arti
(1,0)	Jenis lantai	1.2166	Kecenderungan rumah tangga miskin penerima Raskin saja adalah 1.2166 kali dibandingkan dengan rumah tangga miskin bukan penerima Raskin dan bukan penerima KIS.
(0,1)	Jenis lantai	0.2014	Kecenderungan rumah tangga miskin penerima KIS saja adalah 0.2014 kali dibandingkan dengan rumah tangga miskin bukan penerima Raskin dan bukan penerima KIS.
(1,1)	Jenis lantai	26.0183	Kecenderungan rumah tangga miskin penerima Raskin dan penerima KIS adalah 26.0183 kali dibandingkan dengan rumah tangga miskin bukan penerima Raskin dan bukan penerima KIS.

#### 4. KESIMPULAN

Dari analisis yang telah dilakukan diperoleh hasil yang dapat disimpulkan yaitu :

- 1) Dari 14 indikator rumah tangga miskin yang ditetapkan oleh BPS, hanya 4 indikator yang valid dan dapat digunakan sebagai variabel prediktor dalam analisis penerima Raskin dan KIS dengan menggunakan regresi logistik biner bivariat di Desa Talang Pauh yaitu jenis lantai tempat tinggal, jenis dinding, tindakan pengobatan jika sakit, dan pendidikan terakhir kepala rumah tangga.
- 2) Dari 4 variabel prediktor diperoleh satu variabel prediktor yang signifikan yaitu jenis lantai tempat tinggal.
- 3) Model regresi logistik biner bivariat :

$$g_1(x) = \ln \left[ \frac{\pi_1(x)}{1 - \pi_1(x)} \right] = 3.0804 + 0.1961 X_2$$

$$g_2(x) = \ln \left[ \frac{\pi_2(x)}{1 - \pi_2(x)} \right] = 1.3964 - 1.6021 X_2$$

## 5. DAFTAR PUSTAKA

- [1] Agresti, A. 2007. *An Introduction to Categorical Data Analysis*. Hoboken, New Jersey : John wiley & Sons.
- [2] Anonim. 2008. *Analisis dan Perhitungan Tingkat Kemiskinan*. Badan Pusat Statistik. Jakarta, Indonesia.
- [3] Cessie, L. and J.C. Houwelingen. 1994. *Logistic Regression for Correlated Binary Data*. *Applied Statistics*, 42, hal. 95-108.
- [4] Hosmer, D.W. and S. Lemeshow. 2000. *Applied Logistic Regression Second Ed*. New York – Singapore : John wiley & Sons. Inc.
- [5] Johnson, R.A. and D.W. Winchern. 2007 *Applied Multivariate Statistical Analysis*. Prentice Hall. United Stated.
- [6] McCullagh, P. and J.A. Nelder. 1989. *Generalized Linear Models, Second Edition*. Chapman and Hall, London.
- [7] Sayyida dan Zain, I. 2011. *Analisis Partisipasi Ekonomi Perempuan dengan Metode Regresi Logistik Biner Bivariat*. Fakultas Matematika dan Ilmu Pengetahuan Alam. Institut Teknologi Sepuluh Nopember, Surabaya. Diakses pada tanggal 12 April 2016. <http://digilib.its.ac.id/public/ITS-Master-17978-1309201002-paperpdf.pdf>
- [8] Spring. 2009. *Modelling of African Farm Dinamic Using Bivariate Binary Logistic Regression in WinBUGs*. Carolina : LUND University.
- [9] Viestri, D.A. dan Puhadi. 2015. *Pemodelan Log Linier dan regresi Logistik Biner Bivariat pada hasil Medical Check-Up pegawai negeri Sipil (PNS) Institut Teknologi Sepuluh Nopember (ITS)*. *Jurnal Sains dan Seni ITS*. Vol. 4 hal. 2337-3520.
- [10] Wahyudi, C.D., I.N.Latra, , dan V. Ratnasari. 2013. *Pemodelan Kemiskinan Perdesaan dan Perkotaan dengan Pendekatan Garis Kemiskinan Menggunakan Model Probit Biner Bivariat di Propinsi Bengkulu*. Seminar Nasional Sains dan Teknologi V, Lembaga Penelitian Universitas Lampung. Diakses pada tanggal 12 April 2016. <http://satek.unila.ac.id/wp-content/uploads/2014/03/1-138.pdf>

ANALISIS RANTAI MARKOV UNTUK MENGETAHUI PELUANG  
PERPINDAHAN MEREK KARTU INTERNET GSM PRABAYAR  
(Studi Kasus: Program Studi Matematika Fakultas Matematika dan  
Ilmu Pengetahuan Alam)

Yuli Inda Pertiwi<sup>1</sup>, Sigit Nugroho<sup>2</sup>, dan Dian Agustina<sup>2</sup>

<sup>1</sup>Mahasiswa Jurusan Statistika, Prodi Matematika, FMIPA, UNIB, Bengkulu

<sup>2</sup>Dosen Pembimbing Jurusan Statistika, Prodi Matematika, FMIPA, UNIB, Bengkulu  
[Yuli.ind14@gmail.com](mailto:Yuli.ind14@gmail.com)

**ABSTRAK**

Penelitian ini bertujuan untuk mengetahui peluang perpindahan merek kartu internet GSM ke merek kartu internet GSM lainnya di kalangan mahasiswa Matematika FMIPA UNIB dengan menggunakan analisis rantai markov. Data yang digunakan dalam penelitian ini adalah data primer yang dikumpulkan melalui penyebaran kuesioner pada mahasiswa Matematika FMIPA UNIB. Total sampel dalam penelitian ini adalah 159 responden. Berdasarkan hasil penelitian diketahui bahwa merek kartu internet Hutchison (3) merupakan merek yang mampu mempertahankan jumlah konsumennya sebesar 77,11% dan mempunyai peluang terbesar dalam memperoleh konsumen sehingga pada periode berikutnya merek kartu internet Hutchison (3) memiliki konsumen sebesar 67,72%, sedangkan merek kartu internet XL Axiata merupakan merek yang mempunyai peluang terkecil dalam mempertahankan konsumennya yaitu 0% dan pada periode berikutnya XL Axiata memiliki konsumen sebesar 1,21%. Periode seimbang untuk setiap merek kartu internet GSM terjadi pada periode ke 16.

**Kata Kunci:** Analisis Rantai Markov, Produk Kartu Internet GSM Prabayar

## I. PENDAHULUAN

Perkembangan teknologi yang pesat saat ini berdampak pada semua bidang kehidupan, termasuk bidang komunikasi. Salah satu jenis sarana telekomunikasi yang banyak dipilih masyarakat dengan alasan praktis dan mudah dibawa ke mana saja adalah telepon seluler (*handphone*). Berbicara mengenai telepon seluler maka tak lepas dari operator seluler sebagai penyedia jasa dari telepon seluler. Jasa dari operator seluler sendiri bisa didapat jika menggunakan produk dari operator seluler tersebut yaitu kartu seluler (Sri dan Nuraeni, 2014).

Kelahiran industri seluler di Indonesia di dominasi oleh dua operator seluler berbasis GSM (*Global System for Mobile Communication*) yaitu PT. Telkomsel (Telkomunikasi Seluler Indonesia) dan PT. Satelindo (Satelit Palapa Indonesia) atau yang sekarang dikenal dengan Indosat Satelindo. Beberapa tahun kemudian hadir operator seluler lainnya seperti PT Excelcomindo Pratama (Excelcom, sekarang XL Axiata) dan PT. Hutchison. Di Indonesia pada tahun 2009, telah beroperasi sejumlah 10 operator dengan perkiraan jumlah pelanggan sekitar 175,18 juta. Sebagian besar operator telah meluncurkan

layanan 3G dan 3,5G. Jumlah pengguna seluler di Indonesia bulan Juni 2010 diperkirakan mencapai hingga 180 juta pelanggan, atau mencapai sekitar 80 persen populasi penduduk (Anonim, 2015).

Sejalan dengan semakin banyaknya produk kartu seluler di pasaran, konsumen dapat memilih berbagai pilihan merek sesuai yang dikehendakinya. Hal tersebut tentunya dapat mengakibatkan konsumen berganti operator seluler setiap saat, dengan terjadinya pergantian operator seluler maka dengan pasti konsumen melakukan perpindahan merek (*brand switching*) (Sri dan Nuraeni, 2014).

Definisi dari *brand switching* adalah perpindahan merek yang dilakukan oleh pelanggan untuk setiap waktu penggunaan. *Brand switching* ditandai dengan adanya perbedaan yang signifikan antar merek, dalam hal ini konsumen tidak mengetahui banyak tentang katagori produk yang ada. Hal ini harus dicermati dengan baik oleh para pemasar agar perusahaan dapat mempertahankan dan meningkatkan pasar guna mengungguli para pesaing dengan produk atau merek yang ditawarkan (Ribhan, 2006).

Persaingan ketat antar perusahaan kartu seluler telah menjadikan perusahaan kartu seluler berlomba-lomba memberikan tawaran yang menarik kepada pelanggan maupun calon pelanggannya, dengan tujuan untuk menarik pelanggan sebanyak-banyaknya. Mengingat banyaknya pilihan merek kartu seluler yang ditawarkan di pasaran, serta sering berubahnya selera konsumen maka tidak jarang dalam kurun waktu singkat seseorang pengguna berganti merek dari satu merek ke merek lainnya.

Perhitungan peluang perpindahan merek dapat dihitung dengan menggunakan metode rantai markov. Menurut Langi (2011), rantai markov merupakan sebuah proses stokastik, dimana kejadian pada masa mendatang hanya bergantung pada kejadian hari ini dan tidak bergantung pada kejadian masa lampau. Rantai

markov dalam proses stokastik seringkali digunakan dalam penyelesaian masalah industri khususnya dalam masalah penentuan pangsa pasar (*market share*). Dalam hal ini parameter waktu digunakan untuk menentukan perubahan ataupun besarnya *market share* yang didapat oleh sebuah produk (Halim, 2002). Kemungkinan perpindahan merek oleh konsumen dalam jumlah besar dan berakibat pada *market share* mengharuskan perusahaan untuk mengetahui dan mengantisipasi perubahan *market share* tersebut. Oleh karena itu penelitian ini bertujuan untuk mengetahui peluang perpindahan merek kartu internet GSM ke merek kartu internet GSM lainnya di kalangan mahasiswa Matematika FMIPA UNIB.

## II. LANDASAN TEORI

### A. GSM (Global System for Mobile Communication)

*Global System for Mobile Communication* (GSM mulanya singkatan dari *Groupe Special Mobile*) adalah sebuah teknologi komunikasi seluler yang bersifat digital. Teknologi ini memanfaatkan gelombang mikro dan pengiriman sinyal yang dibagi berdasarkan waktu, sehingga sinyal informasi yang dikirim akan sampai pada tujuan. GSM dijadikan standar global untuk komunikasi seluler sekaligus sebagai teknologi seluler yang paling banyak digunakan orang di seluruh dunia (Anonim, 2015).

### B. Konsep Merek

Menurut Rangkuti (2004) dalam Ribhan (2006), *American Marketing Association* mendefinisikan merek sebagai nama, istilah, tanda, simbol dan rancangan atau kombinasi untuk mengidentifikasi barang dan jasa dari seorang atau sekelompok penjual dan membedakan dari produk pesaing. Tujuan pemberian merek adalah untuk mengidentifikasi produk atau jasa yang ditawarkan pesaing, pengertiannya adalah:

- a. *Brand name* (nama merek) yang merupakan bagian dari merek yang dapat diucapkan.
- b. *Brand mark* (tanda merek) yang merupakan bagian dari merek yang dapat dikenali namun tidak dapat diucapkan seperti lambing, desain huruf atau nama khusus.
- c. *Trade mark* (tanda merek dagang) yang merupakan merek atau bagian dari merek yang dilindungi hukum karena

kemampuannya untuk menghasilkan sesuatu yang istimewa.

- d. *Copyright* (hak cipta) merupakan hak istimewa yang dilindungi oleh undang-undang, menerbitkan, dan menjual karya tulis, karya musik, karya seni.

### C. Perpindahan Merek

*Brand switching* adalah perpindahan atau peralihan merek yang digunakan oleh konsumen untuk setiap waktu penggunaan produk. Dalam menentukan pilihannya konsumen selalu mempertimbangkan nilai maksimal yang akan mereka dapatkan dengan membeli suatu produk tertentu (Prastya, 2013). Perpindahan merek (*brand switching*) adalah saat dimana seorang pelanggan atau sekelompok pelanggan berpindah kesetiaan dari satu merek sebuah produk tertentu ke merek produk lainnya. Perpindahan merek (*brand switching*) merupakan fenomena yang sering terjadi pada berbagai pasar, terutama pasar persaingan sempurna dimana terdapat berbagai macam produk sejenis dengan harga yang bersaing sehingga memudahkan konsumen melakukan pembelian bervariasi (*variety seeking*) (Debora dan Yoestini, 2012).

### D. Proses Stokastik

Proses stokastik  $\{X_n\}$  adalah kumpulan dari variabel acak. Indeks  $n$  diartikan sebagai waktu dan  $X_n$  adalah *state* pada waktu  $n$ . Himpunan  $n$  disebut ruang parameter atau himpunan indeks proses. Jika nilai dari himpunan  $n$  dapat dihitung maka proses stokastik dikatakan proses waktu diskrit. Misalnya,  $\{X_n, n = 0, 1, 2, \dots\}$  adalah proses stokastik waktu diskrit dengan indeks bilangan

bulat tak negatif sedangkan  $\{X_n, n \geq 0\}$  adalah proses stokastik waktu kontinu dengan indeks bilangan *real* tak negatif (Ross, 2010).

### E. Rantai Markov (Markov Chains)

Teori rantai markov pertama kali ditemukan oleh Andrey Andreyevich Markov, matematikawan terkenal dari Rusia. Markov terkenal lewat teori bahwa suatu kejadian berikutnya tergantung hanya pada keadaan saat ini dan bukan pada kejadian masa lalu (Lim dan Sarjono, 2012).

Proses stokastik  $\{X_n\}$  dikatakan memiliki sifat markov jika untuk  $n = 0, 1, 2, \dots$ . Jika  $X_n = i$  maka proses dikatakan berada dalam *state*  $i$  pada waktu  $n$  dan setiap barisan  $i, j, i_0, i_1, \dots, i_{n-1}$  berlaku:

$$\begin{aligned} P\{X_{n+1} = j | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i\} \\ = P\{X_{n+1} = j | X_n = i\} \end{aligned} \quad (2.1)$$

Dengan kata lain, sifat markov adalah peluang bersyarat dari setiap kejadian masa depan yang diberikan kejadian masa lalu, dan keadaan sekarang  $X_n = i$  tidak bergantung dari kejadian masa lalu dan hanya bergantung pada keadaan sekarang (Hillier dan Lieberman, 2001).

Menurut Langi (2011), rantai markov terdefinisi oleh matriks peluang transisi. Matriks peluang transisi adalah suatu matriks yang membuat informasi yang mengatur perpindahan sistem dari suatu *state* ke *state* lainnya. Matriks peluang transisi  $P_{ij}$  adalah tetap dan tidak bergantung pada waktu  $n$ , dimana  $P_{ij}$  adalah peluang transisi satu langkah yang bergerak dari *state*  $i$  ke *state*  $j$ .

#### 1. Peluang Transisi Satu Langkah

Jika sebuah rantai markov  $P\{X_{n+1} = j | X_n = i\} = P_{ij}$  dengan ruang *state*  $\{n = 0, 1, 2, \dots\}$ , maka nilai  $P_{ij}$  merupakan peluang bahwa ketika berada di dalam *state*  $i$ , selanjutnya proses akan melakukan transisi ke *state*  $j$  (Ross, 2010). Misalkan  $P$  menyatakan matriks peluang transisi  $m \times m$ . Matriks peluang transisi  $P$  dapat ditulis sebagai

$$P = \begin{matrix} & \text{State} & 0 & 1 & \dots & m \\ \begin{matrix} 0 \\ 1 \\ \vdots \\ m \end{matrix} & \begin{bmatrix} P_{00} & P_{01} & \dots & P_{0m} \\ P_{10} & P_{11} & \dots & P_{1m} \\ \vdots & \vdots & \vdots & \vdots \\ P_{m0} & P_{m1} & \dots & P_{mm} \end{bmatrix} \end{matrix} \quad (2.2)$$

Sistem dalam *state*  $i$  adalah waktu  $n$ , jadi sistem akan melakukan proses ke suatu *state*  $j$  pada waktu  $n + 1$ . Ini berarti bahwa untuk setiap  $i$ ,

$$\begin{aligned} \sum_{j=1}^m P\{X_{n+1} = j | X_n = i\} = 1 \\ \sum_{j=1}^m P_{ij} = 1 \end{aligned} \quad (2.3)$$

setiap anggota dalam matriks  $P$  harus positif. Oleh karena itu, semua anggota dalam matriks peluang transisi adalah tidak negatif, dan anggota dalam setiap baris harus berjumlah 1 (Winston, 2004).

#### 2. Peluang Transisi $n$ -langkah

Ross (2010) mendefinisikan peluang transisi  $n$ -langkah  $P_{ij}^n$  sebagai peluang bahwa proses pada *state*  $i$  akan berada pada *state*  $j$  setelah proses mengalami  $n$  tambahan transisi. Jadi,

$$P_{ij}^n = P\{X_{n+k} = j | X_k = i\}, \quad n, i, j \geq 0 \quad (2.4)$$

$P_{ij}^n$  disebut peluang transisi  $n$ -langkah dari *state*  $i$  ke *state*  $j$ .

Menurut Hillier dan Lieberman (2001), peluang transisi  $n$ -langkah  $P_{ij}^n$  adalah peluang bersyarat bahwa sistem akan berada pada *state*  $j$  setelah mengalami proses  $n$ -langkah, proses dimulai pada *state*  $i$  pada waktu  $t$ . Matriks peluang transisi  $n$ -langkah  $P_{ij}^n$  dapat ditulis sebagai

$$P^n = \begin{matrix} & \text{State} & 0 & 1 & \dots & m \\ \begin{matrix} 0 \\ 1 \\ \vdots \\ m \end{matrix} & \begin{bmatrix} P_{00}^n & P_{01}^n & \dots & P_{0m}^n \\ P_{10}^n & P_{11}^n & \dots & P_{1m}^n \\ \dots & \dots & \dots & \dots \\ P_{m0}^n & P_{m1}^n & \dots & P_{mm}^n \end{bmatrix} \end{matrix} \quad (2.4)$$

Ketika  $n = 1$ , diketahui bahwa  $P_{ij}^1 = P_{ij}$ . Karena  $P_{ij}^n$  adalah peluang bersyarat, peluang tidak negatif, dan karena proses harus melakukan transisi ke beberapa *state*, sehingga:

$$P_{ij}^n \geq 0 \text{ untuk semua } i \text{ dan } j; n = 0, 1, 2 \dots$$

dan

$$\sum_{j=1}^m P_{ij}^n = 1 \text{ untuk semua } i; n = 0, 1, 2 \dots$$

### F. Persamaan Chapman-Kolmogorov

Persamaan Chapman-Kolmogorov merupakan suatu metode untuk menghitung peluang transisi  $n$ -langkah:

$$\begin{aligned} P_{ij}^{n+m} = \\ P\{X_{n+m} = j | X_0 = i\} \\ \text{untuk semua } n, m, i, j \geq 0 \end{aligned}$$

$$= \sum_{k=0}^{\infty} P_{kj}^m P_{ik}^n \quad (2.5)$$

Persamaan (2.5) menunjukkan bahwa  $P_{ik}^n P_{kj}^m$  adalah peluang dengan titik awal dari *state*  $i$ , dalam perubahan dari *state*  $i$  ke *state*  $j$  sebanyak  $n + m$  langkah, proses akan berada

dalam *state k* setelah tepat  $n$ -langkah. Oleh karena itu, dapat disimpulkan dari semua *state k* menghasilkan peluang bahwa proses akan berada di *state j* setelah  $n + m$  langkah.

Misalkan  $\mathbf{P}^n$  menunjukkan matriks peluang transisi  $n$ -langkah  $P_{ij}^n$ , maka persamaan (2.5) menyatakan bahwa

$$\mathbf{P}^{n+m} = \mathbf{P}^n \cdot \mathbf{P}^m$$

dengan demikian, matriks peluang transisi  $n$ -langkah  $\mathbf{P}^n$  dapat diperoleh dengan mengalikan  $n$  dari matriks peluang transisi satu langkah  $\mathbf{P}$  (Ross, 2010).

## G. Klasifikasi State

Peluang transisi yang terkait dengan *state* memainkan peran penting dalam studi rantai Markov. Untuk mempelajari sifat rantai Markov, berikut beberapa konsep dan definisi tentang *state*.

### 1. Accessible

*State j* dikatakan dapat dicapai (*accessible*) dari *state i* jika terdapat bilangan bulat  $n \geq 0$  sehingga  $P_{ij}^n > 0$ . Dengan demikian, *state j* yang dapat dicapai (*accessible*) dari *state i* artinya adalah ketika sistem dimulai dari *state i* akhirnya sistem akan pindah ke *state j* (Hillier dan Lieberman, 2001).

### 2. Communicate

Menurut Ross (2010), jika *state j* dicapai dari *state i* dinotasikan dengan  $i \rightarrow j$  dan *state i* dapat dicapai dari *state j* dinotasikan dengan  $j \rightarrow i$ , maka *state i* dan *j* dikatakan saling berkomunikasi (*communicate*) dinotasikan dengan  $i \leftrightarrow j$ .

Secara umum hubungan komunikasi memenuhi tiga sifat berikut (Hillier dan Lieberman, 2001):

1. Setiap *state* berkomunikasi dengan dirinya ( $i \leftrightarrow i$ ) karena  $P_{ii}^0 = P\{X_0 = i | X_0 = i\} = 1$ .
2. Jika *state i* berkomunikasi dengan *state j* ( $i \leftrightarrow j$ ), maka *state j* berkomunikasi dengan *state i* ( $j \leftrightarrow i$ ).
3. Jika *state i* berkomunikasi dengan *state j* ( $i \leftrightarrow j$ ) dan *state j* berkomunikasi dengan *state k* ( $j \leftrightarrow k$ ), maka *state i* berkomunikasi dengan *state k* ( $i \leftrightarrow k$ ).

Dari tiga sifat komunikasi diatas, *state* dapat dibagi menjadi satu atau lebih kelas terpisah sehingga *state* berkomunikasi satu sama lain berada dalam kelas yang sama, sebuah kelas dapat terdiri dari satu *state*. Jika hanya ada satu kelas, yaitu semua *state* berkomunikasi maka rantai Markov dikatakan *irreducible* (tidak dapat diperkecil lagi) (Hillier dan Lieberman, 2001).

## 3. Recurrent dan Transient State

Sebuah *state* dikatakan *recurrent state* (berulang) jika setelah memasuki *state i*, proses dengan pasti akan kembali ke *state i* lagi. Oleh karena itu, *state i* dikatakan *recurrent state* (berulang) jika dan hanya jika tidak *transient state*. Sebuah *state* dikatakan *transient state* (sementara) jika setelah memasuki *state i*, proses tidak pernah kembali ke *state i* lagi (Hillier dan Lieberman, 2001).

Ross (2010), untuk proses yang dimulai dari *state i* dan akan kembali ke *state i* didefinisikan sebagai berikut:

$$f_{ii} = \sum_{n=1}^{\infty} f_{ii}^n = f_{ii}^1 + f_{ii}^2 + \dots \quad (2.6)$$

$f_{ii}$  menunjukkan peluang bahwa, proses yang mulai dari *state i* akan masuk kembali ke *state i*. *State i* dikatakan *recurrent* (berulang) jika  $f_{ii} = 1$  dan *transient* (sementara) jika  $f_{ii} < 1$ .

**Proposition 2.1** (Ross, 2010:206)

*State i* adalah:

Recurrent jika

$$\sum_{n=1}^{\infty} P_{ii}^n = \infty \quad (2.7)$$

Transient jika

$$\sum_{n=1}^{\infty} P_{ii}^n < \infty \quad (28)$$

## 4. Absorbing State

*State i* dikatakan *absorbing* (menyerap) jika  $P_{ii} = 1$  (Winston, 2004). Menurut Hillier dan Lieberman (2001), sebuah *state* dikatakan *absorbing* jika setelah memasuki *state i*, proses tidak akan pernah meninggalkan *state i* ini lagi. Artinya tidak ada *state* yang dapat dicapai dari *state* tersebut.

## 5. Sifat Periode

*State i* dikatakan memiliki periode  $d(i)$  jika  $P_{ii}^n = 0$ , dimana  $n$  tidak habis dibagi  $d$ , dan  $d$  adalah bilangan bulat terbesar. Periode  $d(i)$  merupakan FPB (faktor persekutuan terbesar) atau gcd (*greatest common divisor*) dari seluruh  $n = 1, 2, \dots$  dimana  $P_{ii}^n > 0$ .

$$d(i) = \gcd\{n \geq 1 | P_{ii}^n > 0\} \quad (2.9)$$

- Jika  $d(i) = 1$ , maka *state i* disebut aperiodik

- Jika  $d(i) > 1$ , maka *state i* disebut periodik

Jika *state i* memiliki periode  $d$ , dan *state i* dan *state j* berkomunikasi, maka *state j* juga memiliki periode  $d$  (Ross, 2010).

## H. Limiting Probability

State  $i$  yang *recurrent* (berulang) dikatakan positif *recurrent* jika proses dimulai pada state  $i$ , maka waktu yang diharapkan proses untuk sampai kembali ke state  $i$  terbatas. Sebaliknya terdapat pula *states recurrent* (berulang) yang tidak positif *recurrent*, *states* ini disebut *null recurrent*. Untuk suatu rantai markov, semua *state recurrent* diklasifikasikan menjadi keadaan positif *recurrent* atau *null recurrent* dengan menyatakan rata-rata waktu *recurrent* (*mean recurrent time*)

$$\mu_i = \sum_{n=1}^{\infty} n \cdot f_{ii}^n \quad (2.10)$$

- Jika  $\mu_i < \infty$ , maka state  $i$  disebut positif *recurrent*

- Jika  $\mu_i = \infty$ , maka state  $i$  disebut *null recurrent*

untuk setiap state  $i$ ,  $f_{ii}$  menunjukkan peluang bahwa proses yang mulai dari state  $i$  akan masuk kembali ke state  $i$ . State yang positif *recurrent* dan aperiodek disebut *ergodic*.

**Teorema 2.1** ( Ross, 2010:215)

## III. METODE PENELITIAN

Metode penelitian yang digunakan adalah studi kasus atau penelitian terapan dengan metode survei. Data yang digunakan dalam penelitian ini adalah data primer yang diperoleh secara langsung berupa jawaban responden yang diperoleh dari pengisian kuesioner oleh mahasiswa S1 Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Bengkulu. Teknik analisis data yang dilakukan terdiri dari beberapa tahapan sebagai berikut:

1. Melakukan penyebaran kuesioner.
2. Input data yang diperoleh dari kuesioner.
3. Membuat tabel jumlah pengguna kartu internet GSM Prabayar dari masing-masing merek.
4. Membuat tabel *brand switching* yaitu data perubahan atau pergeseran dari satu merek ke merek lainnya.
5. Menentukan peluang dari setiap keadaan.
6. Membuat matriks peluang transisi ( $P$ ).
7. Memprediksi pangsa pasar menggunakan matriks peluang transisi  $n$ -langkah yaitu mengalikan matriks peluang transisi dengan dirinya sendiri sebanyak  $n$ -langkah sampai mencapai keadaan yang seimbang, yaitu keadaan dimana persaingan pasar mengalami keadaan yang tetap.
8. Menentukan klasifikasi state rantai markov.
9. Menentukan *limiting probability*.

## IV. Hasil dan Pembahasan

Jika untuk setiap rantai markov yang *irreducible ergodic*  $\lim_{n \rightarrow \infty} P_{ij}^n$  ada dan independen pada state  $i$ , maka

$$\lim_{n \rightarrow \infty} P_{ij}^n = \pi_j > 0, \quad i, j \geq 0 \quad (2.11)$$

dimana  $\pi_j$  merupakan solusi nonnegatif dari:

$$\pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij}$$

$$\sum_{j=0}^{\infty} \pi_j = 1 \quad (2.12)$$

Menurut Hillier dan Lieberman (2001)  $\pi = [\pi_0, \pi_1, \pi_2, \dots, \pi_n]$  disebut distribusi stasioner dari rantai markov. Artinya jika state awal berada di state  $j$  diberikan oleh  $\pi_j$ , maka peluang untuk menentukan proses state  $j$  pada waktu  $n = 1, 2, \dots$  juga diberikan oleh  $\pi_j$ , yaitu jika

$$P\{X_0 = j\} = \pi_j, \quad j \geq 0$$

maka

$$P\{X_n = j\} = \pi_j, \quad \text{untuk semua } n, j \geq 0$$

## Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data primer yang diperoleh secara langsung dari jawaban responden yang diteliti yaitu berupa jawaban dari pengisian kuesioner oleh mahasiswa S1 Matematika FMIPA UNIB. Pada penelitian ini dibagikan sebanyak 159 kuesioner kepada responden dan semua kuesioner tersebut kembali dalam keadaan baik sehingga dapat digunakan oleh peneliti dalam melakukan pengolahan data.

## Proses Pengolahan Data

### 1. Merek Kartu Internet GSM yang Digunakan

Merek-merek kartu internet GSM yang digunakan oleh responden mahasiswa FMIPA Universitas Bengkulu adalah seperti Tabel 1 berikut:

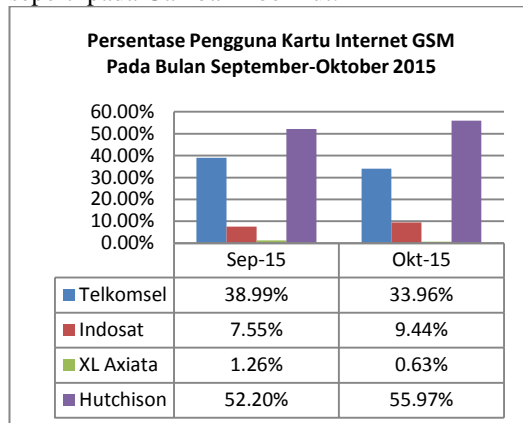
Tabel 1. Merek Kartu Internet GSM dan Jumlah Pengguna Pada Bulan September 2015

No.	Merek Kartu internet GSM	Jumlah responden	Proporsi (%)
1	Telkomsel	62	38,99
2	Indosat	12	7,55
3	XL Axiata	2	1,26
4	Hutchison	83	52,20
Jumlah		159	100

Tabel 1 menunjukkan bahwa 62 responden (38,99%) menggunakan kartu internet Telkomsel. Kemudian 12 responden (7,55%) menggunakan

kartu internet Indosat dan 2 responden (1,26%) menggunakan kartu internet XL Axiata. Selanjutnya kartu internet Hutchison sebesar 83 responden (52,20) merupakan merek kartu internet yang paling banyak digunakan oleh responden.

Pada bulan Oktober 2015 atau periode kedua terjadi perubahan kenaikan dan penurunan persentase pada masing-masing merek kartu internet GSM yang digunakan oleh responden mahasiswa S1 Matematika FMIPA UNIB adalah seperti pada Gambar 1 berikut:



Gambar 1. Persentase Pengguna Merek Kartu Internet GSM Pada Bulan September-Oktober 2015

Tabel 2. Jumlah Perolehan dan Kehilangan Responden Kartu Internet GSM

No	Merek Kartu Internet GSM	Jumlah Responden Sebelumnya	Perolehan	Kehilangan	Jumlah Responden Saat Ini
1	Telkomsel	62	17	25	54
2	Indosat	12	12	9	15
3	XL Axiata	2	1	2	1
4	Hutchison (3)	83	25	19	89

Tabel 2 menunjukkan merek kartu internet Telkomsel saat ini berjumlah 54 responden, sebelumnya pengguna kartu internet Telkomsel berjumlah 62 responden, ini berarti terjadi pengurangan sebanyak 8 responden. Merek kartu internet Telkomsel memperoleh 17 responden dari merek kartu internet GSM lainnya akan tetapi 25 responden berpindah ke merek kartu internet GSM lainnya.

Jumlah pengguna kartu internet Indosat bertambah 3 responden yang sebelumnya 12 responden menjadi 15 responden. Merek kartu internet XL Axiata yang semula digunakan oleh 2 responden sekarang berkurang menjadi 1 responden yang merupakan responden dari merek kartu internet GSM lain yang berpindah ke merek kartu internet XL Axiata karena 2 responden merek

kartu internet XL Axiata yang sebelumnya telah berpindah ke merek kartu internet GSM lainnya.

Merek kartu internet Hutchison adalah merek yang paling banyak digunakan oleh responden baik sekarang maupun sebelumnya. Jumlah pengguna kartu internet Hutchison bertambah 6 responden dari yang sebelumnya 83 responden menjadi 89 responden. Merek kartu internet Hutchison memperoleh 25 responden dari merek kartu internet GSM lainnya akan tetapi 19 responden berpindah ke merek kartu internet GSM lainnya.

## 2. Perpindahan Merek Kartu Internet GSM

Pola perpindahan merek dari satu merek ke merek yang lainnya dapat digambarkan pada Tabel 3 seperti berikut :

Tabel 3. Pola Perpindahan Merek Kartu Internet GSM

Dari Merek	Merek Kartu Internet GSM	Ke Merek				Jumlah Responden Sebelumnya
		Telkomsel	Indosat	XL Axiata	Hutchison	
	Telkomsel	37	7	0	18	62
	Indosat	4	3	0	5	12
	XL Axiata	0	0	0	2	2
	Hutchison	13	5	1	64	83
	Jumlah Responden Saat Ini	54	15	1	89	159

Selanjutnya apabila dari Tabel 3 diasumsikan bahwa pergeseran diantara merek kartu internet dianggap stabil (tetap tidak berubah) maka dapat dibuat Tabel Peluang Transisi ( $P$ ) yang ditunjukkan pada Tabel 4.

Tabel 4. Tabel Peluang Transisi ( $P$ )

Dari Merek	Ke Merek			
	Telkomsel	Indosat	XL Axiata	Hutchison
Telkomsel	0,5968	0,1129	0,0000	0,2903
Indosat	0,3333	0,2500	0,0000	0,4167
XL Axiata	0,0000	0,0000	0,0000	1,0000
Hutchison	0,1566	0,0602	0,0121	0,7711

Pada tabel 7 diperoleh data responden Telkomsel yang tetap menggunakan kartu internet Telkomsel adalah 59,68% yang diperoleh dari  $\frac{37}{62}$ , dengan angka 37 menunjukkan jumlah responden yang tetap menggunakan merek kartu internet Telkomsel dan angka 62 menunjukkan jumlah responden sebelumnya yang berpindah dari merek Telkomsel ke tiga merek lainnya. Kemudian dari merek kartu internet Telkomsel berpindah ke merek kartu internet Indosat 11,29% yang diperoleh dari  $\frac{7}{62}$ , dengan angka 7 menunjukkan jumlah responden yang berpindah dari merek kartu internet Telkomsel ke merek kartu internet Indosat

dan angka 62 menunjukkan jumlah responden sebelumnya yang berpindah dari merek Telkomsel ke tiga merek lainnya. Selanjutnya merek kartu internet Telkomsel berpindah ke merek kartu internet Hutchison yaitu 29,03% yang diperoleh dari  $\frac{18}{62}$ , dengan angka 18 menunjukkan jumlah responden yang berpindah dari merek kartu internet Telkomsel ke merek kartu internet Hutchison dan angka 62 menunjukkan jumlah responden sebelumnya yang berpindah dari merek Telkomsel ke tiga merek lainnya. Begitu seterusnya untuk memperoleh nilai peluang untuk tiga merek kartu internet lainnya.

Misalkan  $P$  merupakan matriks peluang transisi dengan 4 *state* merek kartu internet GSM yang menyatakan peluang perpindahan merek kartu internet GSM di kalangan mahasiswa S1 Matematika FMIPA UNIB, maka dari Tabel 4 dapat dituliskan:

$$P = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \end{matrix} & \begin{bmatrix} 0,5968 & 0,1129 & 0,0000 & 0,2903 \\ 0,3333 & 0,2500 & 0,0000 & 0,4167 \\ 0,0000 & 0,0000 & 0,0000 & 1,0000 \\ 0,1566 & 0,0602 & 0,0121 & 0,7711 \end{bmatrix} \end{matrix}$$

Keterangan:

- 0 = Responden Telkomsel
- 1 = Responden Indosat
- 2 = Responden XL Axiata
- 3 = Responden Hutchison

Anggota pada matriks  $P$  dapat dinotasikan dengan  $P_{ij}$  yang merupakan peluang perpindahan responden dari merek sebelumnya  $i$  ke merek saat ini  $j$ . Nilai peluang setiap  $P_{ij}$  antara 0 dan 1. Sehingga matriks  $P$  disebut juga matriks tak negatif, dan anggota dalam setiap baris adalah 1.

### Prediksi Pangsa Pasar (Market Share) Merek Kartu Internet GSM

Prediksi pangsa pasar merek kartu internet GSM pada periode yang akan datang dapat dilakukan dengan menggunakan matriks peluang transisi  $n$ -langkah yaitu mengalikan matriks peluang transisi dengan dirinya sendiri sebanyak  $n$ -langkah. Berdasarkan hasil perhitungan dengan menggunakan program MATLAB diperoleh hasil bahwa nilai peluang perpindahan dari masing-masing merek kartu internet GSM akan mengalami keadaan seimbang saat langkah ke 16 atau pada bulan Desember 2016, artinya persaingan pasar merek kartu internet GSM dalam hal mempertahankan, memperoleh atau kehilangan konsumen akan mengalami keadaan seimbang pada langkah ke 16. Matriks peluang transisi  $P^{16}$  menunjukkan bahwa kartu internet Hutchison adalah merek yang paling banyak digunakan oleh responden sebesar 59,17% atau sekitar 94 responden. Kemudian Telkomsel di urutan kedua

terbanyak digunakan oleh responden sebesar 30,73% atau 49 responden dan Indosat di urutan berikutnya sebesar 9,38% atau 15 responden. Selanjutnya pada urutan terakhir XL Axiata yaitu 0,72% atau 1 responden. Pada langkah ke 17, 18, dan seterusnya akan diperoleh matriks peluang transisi yang sama karena proses sudah mencapai keadaan tetap.

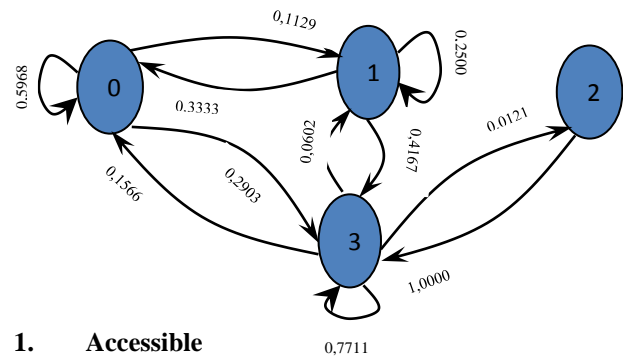
### Menentukan Klasifikasi State Rantai Markov

Klasifikasi *state* rantai markov dapat ditentukan dengan menggunakan matriks peluang transisi  $P$ ,

$$P = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \end{matrix} & \begin{bmatrix} 0,5968 & 0,1129 & 0,0000 & 0,2903 \\ 0,3333 & 0,2500 & 0,0000 & 0,4167 \\ 0,0000 & 0,0000 & 0,0000 & 1,0000 \\ 0,1566 & 0,0602 & 0,0121 & 0,7711 \end{bmatrix} \end{matrix}$$

dengan diagram transisinya sebagai berikut:

Gambar 2. Diagram Transisi



#### 1. Accessible

*State* yang dapat dicapai (*accessible*), yaitu jika terdapat bilangan bulat  $n \geq 0$  sehingga  $P_{ij}^n > 0$ , dari Gambar 2 dapat ditunjukkan *state* yang dapat dicapai (*accessible*) sebagai berikut:

- *State* 0 dapat dicapai dari *state* 1

$$\begin{aligned} P_{01} &= 0,1129 \\ P_{01}^2 &= 0,1131 \\ P_{01}^3 &= 0,1046 \\ &\vdots \end{aligned}$$

artinya ada responden yang setelah menggunakan Telkomsel berpindah ke merek kartu internet Indosat.

- *State* 0 dapat dicapai dari *state* 3

$$\begin{aligned} P_{03} &= 0,2903 \\ P_{03}^2 &= 0,4441 \\ P_{03}^3 &= 0,5206 \\ &\vdots \end{aligned}$$

artinya ada responden yang setelah menggunakan Telkomsel berpindah ke merek kartu internet Hutchison.

- *State* 1 dapat dicapai dari *state* 0

$$\begin{aligned} P_{10} &= 0,3333 \\ P_{10}^2 &= 0,3475 \\ P_{10}^3 &= 0,3309 \end{aligned}$$

⋮

artinya ada responden yang setelah menggunakan Indosat berpindah ke merek kartu internet Telkomsel.

- State 1 dapat dicapai dari state 3

$$P_{13} = 0,4167$$

$$P_{13}^2 = 0,5222$$

$$P_{13}^3 = 0,5608$$

⋮

artinya ada responden yang setelah menggunakan Indosat berpindah ke merek kartu internet Hutchison.

- State 2 dapat dicapai dari state 3

$$P_{23} = 1,0000$$

$$P_{23}^2 = 0,7711$$

$$P_{23}^3 = 0,6772$$

⋮

artinya ada responden yang setelah menggunakan XL Axiata berpindah ke merek kartu internet Hutchison.

- State 3 dapat dicapai dari state 0

$$P_{30} = 0,1566$$

$$P_{30}^2 = 0,2343$$

$$P_{30}^3 = 0,2723$$

⋮

artinya ada responden yang setelah menggunakan Hutchison berpindah ke merek kartu internet Telkomsel.

- State 3 dapat dicapai dari state 1

$$P_{31} = 0,0602$$

$$P_{31}^2 = 0,0792$$

$$P_{31}^3 = 0,0870$$

⋮

artinya ada responden yang setelah menggunakan Hutchison berpindah ke merek kartu internet Indosat.

- State 3 dapat dicapai dari state 2

$$P_{32} = 0,0121$$

$$P_{32}^2 = 0,0093$$

$$P_{32}^3 = 0,0082$$

⋮

artinya ada responden yang setelah menggunakan Hutchison berpindah ke merek kartu internet XL Axiata.

## 2. Communicate

Berdasarkan sifat *communicate*, yaitu jika *state i* berkomunikasi dengan *state j*, maka *state j* berkomunikasi dengan *state i*. Pada Gambar 2 dapat dikatakan bahwa *state* yang saling berkomunikasi adalah

- State 0 saling berkomunikasi dengan state 1. Artinya, ada responden yang setelah menggunakan Telkomsel berpindah ke merek kartu internet Indosat dan ada responden yang setelah menggunakan Indosat berpindah ke merek kartu internet Telkomsel.

- State 0 saling berkomunikasi dengan state 3. Artinya, ada responden yang setelah menggunakan Telkomsel berpindah ke merek kartu internet Hutchison dan ada responden yang setelah menggunakan Hutchison berpindah ke merek kartu internet Telkomsel.

- State 1 dan state 3 dikatakan saling berkomunikasi. Artinya, ada responden yang setelah menggunakan Indosat berpindah ke merek kartu internet Hutchison dan ada responden yang setelah menggunakan Hutchison berpindah ke merek kartu internet Indosat.

- State 2 dan state 3 saling berkomunikasi. Artinya, ada responden yang setelah menggunakan XL Axiata berpindah ke merek kartu internet Hutchison dan ada responden yang setelah menggunakan Hutchison berpindah ke merek kartu internet XL Axiata.

Selanjutnya berdasarkan sifat *communicate*, Jika *state i* berkomunikasi dengan *state j* dan *state j* berkomunikasi dengan *state k*, maka *state i* berkomunikasi dengan *state k*. Gambar 2 menunjukkan bahwa:

- State 0 saling berkomunikasi dengan state 2. Artinya ada responden yang setelah menggunakan Telkomsel berpindah ke merek kartu internet Hutchison dan ada responden yang setelah menggunakan Hutchison berpindah ke merek kartu internet XL Axiata, maka responden yang setelah menggunakan Telkomsel dapat berpindah ke merek kartu internet XL Axiata.

- State 1 saling berkomunikasi dengan state 2. Artinya ada responden yang setelah menggunakan Indosat berpindah ke merek kartu internet Hutchison dan ada responden yang setelah menggunakan Hutchison berpindah ke merek kartu internet XL Axiata, maka responden yang setelah menggunakan Indosat dapat berpindah ke merek kartu internet XL Axiata.

Berdasarkan uraian diatas maka *state 0,1,2* dan *state 3* saling berkomunikasi, jadi hanya ada satu kelas komunikasi yaitu  $\{0,1,2,3\}$  semua *state* berkomunikasi, maka rantai markov dikatakan *irreducible*.

## 3. Recurrent dan Transient State

*State* dikatakan *recurrent* (berulang) jika setelah memasuki *state i*, proses dengan pasti akan kembali ke *state i* lagi sedangkan *state i* dikatakan *transient* (sementara) jika setelah memasuki *state i*, proses tidak pernah kembali ke *state i* lagi. Pada Proposition 2.1 *state* dikatakan *recurrent* (berulang) jika  $\sum_{n=1}^{\infty} P_{ii}^n = \infty$  dan *state i* dikatakan

*transient* (sementara) jika  $\sum_{n=1}^{\infty} P_{ii}^n < \infty$ , dari Gambar 2 dapat ditunjukkan *recurrent* dan *transient state* sebagai berikut:

- *State 0*

$$\begin{aligned}
 P_{00} &= 0,5968 \\
 P_{00}^2 &= P_{00}P_{00} + P_{01}P_{10} + P_{03}P_{30} \\
 &= 0,4393 \\
 P_{00}^3 &= P_{00}P_{00}P_{00} + P_{00}P_{01}P_{10} + \dots + \\
 &P_{03}P_{33}P_{30} \\
 &= 0,3694 \\
 &\vdots \\
 \sum_{n=1}^{\infty} P_{00}^n &= P_{00}^1 + P_{00}^2 + P_{00}^3 + \dots \\
 &= 0,5986 + 0,4393 + \\
 &0,3694 + \dots \\
 &= \infty
 \end{aligned}$$

*State*  $P_{00} = \infty$ , maka *state 0* dikatakan *recurrent*, artinya ada responden yang setelah menggunakan kartu internet Telkomsel pasti kembali menggunakan kartu internet Telkomsel.

- *State 1*

$$\begin{aligned}
 P_{11} &= 0,2500 \\
 P_{11}^2 &= P_{10}P_{01} + P_{11}P_{11} + P_{13}P_{31} \\
 &= 0,1225 \\
 P_{11}^3 &= P_{10}P_{00}P_{01} + P_{10}P_{01}P_{11} + \dots + \\
 &P_{13}P_{33}P_{31} \\
 &= 0,1020 \\
 &\vdots \\
 \sum_{n=1}^{\infty} P_{11}^n &= P_{11}^1 + P_{11}^2 + P_{11}^3 + \dots \\
 &= 0,2500 + 0,1252 + \\
 &0,1020 + \dots \\
 &= \infty
 \end{aligned}$$

*State*  $P_{11} = \infty$ , maka *state 1* dikatakan *recurrent*, artinya ada responden yang setelah menggunakan kartu internet Indosat pasti kembali menggunakan kartu internet Indosat.

- *State 2*

$$\begin{aligned}
 P_{22} &= 0,0000 \\
 P_{22}^2 &= P_{23}P_{32} \\
 &= 0,0121 \\
 P_{22}^3 &= P_{23}P_{33}P_{32} \\
 &= 0,0093 \\
 &\vdots \\
 \sum_{n=1}^{\infty} P_{22}^n &= P_{22}^1 + P_{22}^2 + P_{22}^3 + \dots \\
 &= 0,0000 + 0,0121 + \\
 &0,0093 + \dots \\
 &= \infty
 \end{aligned}$$

*State*  $P_{22} = \infty$ , maka *state 2* dikatakan *recurrent*, artinya ada responden yang setelah menggunakan kartu internet XL Axiata pasti kembali menggunakan kartu internet XL Axiata.

- *State 3*

$$\begin{aligned}
 P_{33} &= 0,7711 \\
 P_{33}^2 &= P_{30}P_{03} + P_{31}P_{13} + P_{32}P_{23} \\
 &\quad + P_{33}P_{33} \\
 &= 6772 \\
 P_{33}^3 &= P_{30}P_{00}P_{03} + P_{30}P_{01}P_{13} + \dots + \\
 &P_{33}P_{33}P_{33} \\
 &= 0,6325 \\
 &\vdots \\
 \sum_{n=1}^{\infty} P_{33}^n &= P + P_{33}^2 + P_{33}^3 + \dots \\
 &= 0,7711 + 0,6772 + \\
 &0,6325 + \dots \\
 &= \infty
 \end{aligned}$$

*State*  $P_{33} = \infty$ , maka *state 3* dikatakan *recurrent*, artinya ada responden yang setelah menggunakan kartu internet Hutchison pasti kembali menggunakan kartu internet Hutchison.

Dari uraian di atas maka dapat dikatakan bahwa *state 0,1,2*, dan *state 3* merupakan *recurrent state* karena setelah memasuki *state i*, proses dengan pasti akan kembali ke *state i* lagi dan  $\sum_{n=1}^{\infty} P_{ii}^n = \infty$ .

#### 4. Absorbing State

Berdasarkan matriks peluang transisi  $P$  dapat dilihat bahwa tidak ada *state* yang *absorbing*, karena  $P_{ii} \neq 1$ . Artinya jika setelah memasuki *state i*, proses akan meninggalkan *state i*. Pada Gambar 2 dapat dikatakan bahwa:

- *State 0* tidak *absorbing*, karena  $P_{00} = 0,5968$  artinya ada responden yang setelah menggunakan kartu internet Telkomsel berpindah ke merek kartu internet lainnya.
- *State 1* tidak *absorbing*, karena  $P_{11} = 0,25000$  artinya ada responden yang setelah menggunakan kartu internet Indosat berpindah ke merek kartu internet lainnya.
- *State 2* tidak *absorbing*, karena  $P_{22} = 0,0000$  artinya ada responden yang setelah menggunakan kartu internet XL Axiata berpindah ke merek kartu internet lainnya.
- *State 3* tidak *absorbing*, karena  $P_{33} = 0,7711$  artinya ada responden yang setelah menggunakan kartu internet Hutchison berpindah ke merek kartu internet lainnya.

#### 5. Sifat Periodeitas

Persamaan (2.9) dapat digunakan untuk menentukan *state* yang memiliki periode  $d(i)$  sebagai berikut:

- *State 0*

$$\begin{aligned}
 n &\geq 1, P_{00}^n > 0 \\
 n = 2 &\rightarrow P_{00}^2 = 0,4393 > 0 \\
 n = 3 &\rightarrow P_{00}^3 = 0,3694 > 0 \\
 n = 4 &\rightarrow P_{00}^4 = 0,3368 > 0 \\
 d(0) &= \gcd\{n \geq 1, P_{00}^n > 0\} \\
 &= \gcd\{2,3,4, \dots\} \\
 &= 1
 \end{aligned}$$

Periode *state* 0 yaitu  $d(0) = 1$  artinya *state* 0 dikatakan aperiodik.

- *State* 1  $n \geq 1, P_{11}^n > 0$   
 $n = 2 \rightarrow P_{11}^2 = 0,1252 > 0$   
 $n = 3 \rightarrow P_{11}^3 = 0,1020 > 0$   
 $n = 4 \rightarrow P_{11}^4 = 0,0966 > 0$   
 $d(1) = \gcd\{n \geq 1, P_{11}^n > 0\}$   
 $= \gcd\{2,3,4, \dots\}$   
 $= 1$

Periode *state* 1 yaitu  $d(1) = 1$  artinya *state* 1 dikatakan aperiodik.

- *State* 2  $n \geq 1, P_{22}^n > 0$   
 $n = 2 \rightarrow P_{22}^2 = 0,0121 > 0$   
 $n = 3 \rightarrow P_{22}^3 = 0,0093 > 0$   
 $n = 4 \rightarrow P_{22}^4 = 0,0082 > 0$   
 $d(0) = \gcd\{n \geq 1, P_{22}^n > 0\}$   
 $= \gcd\{2,3,4, \dots\}$   
 $= 1$

Periode *state* 2 yaitu  $d(2) = 1$  artinya *state* 2 dikatakan aperiodik.

- *State* 3  $n \geq 1, P_{33}^n > 0$   
 $n = 2 \rightarrow P_{33}^2 = 0,6772 > 0$   
 $n = 3 \rightarrow P_{33}^3 = 0,6325 > 0$   
 $n = 4 \rightarrow P_{33}^4 = 0,6112 > 0$   
 $d(0) = \gcd\{n \geq 1, P_{33}^n > 0\}$   
 $= \gcd\{2,3,4, \dots\}$   
 $= 1$

Periode *state* 3 yaitu  $d(3) = 1$  artinya *state* 3 dikatakan aperiodik.

Dari uraian di atas maka dapat dikatakan bahwa periode *state* 0,1,2, dan *state* 3 merupakan aperiodik karena  $d(0) = d(1) = d(2) = d(3) = 1$ . Artinya untuk setiap periode (bulan) akan ada responden yang tetap menggunakan merek kartu internet yang digunakan sebelumnya.

### Menentukan Limiting Probability

Berdasarkan Teorema 2.1 *limiting probability* dapat ditentukan dengan menggunakan matriks peluang transisi  $P$  dan menentukan distribusi stasioner  $[\pi_0, \pi_1, \pi_2, \pi_3]$  sebagai berikut:  
 $[\pi_0 \ \pi_1 \ \pi_2 \ \pi_3]$

$$= [\pi_0 \ \pi_1 \ \pi_2 \ \pi_3] \begin{bmatrix} 0,5968 & 0,1129 & 0,0000 & 0,2903 \\ 0,3333 & 0,2500 & 0,0000 & 0,4167 \\ 0,0000 & 0,0000 & 0,0000 & 1,0000 \\ 0,1566 & 0,0602 & 0,0121 & 0,7711 \end{bmatrix}$$

$$\begin{aligned} \pi_0 &= 0,5968\pi_0 + 0,3333\pi_1 + 0,1566\pi_3 \\ \pi_1 &= 0,1129\pi_0 + 0,2500\pi_1 + 0,0602\pi_3 \\ \pi_2 &= 0,0121\pi_3 \\ \pi_3 &= 0,2903\pi_0 + 0,4167\pi_1 + \pi_2 \\ &\quad + 0,7711\pi_3 \end{aligned}$$

dengan melakukan substitusi dan eliminasi, maka didapatkan:

$$\begin{aligned} \pi_0 &= 0,3073, & \pi_1 &= 0,0938, \\ \pi_2 &= 0,0072, & \pi_3 &= 0,5917 \end{aligned}$$

Artinya untuk jangka panjang peluang responden mahasiswa S1 Matematika FMIPA UNIB akan tetap menggunakan merek kartu internet Telkomsel adalah 30,73%, peluang responden akan tetap menggunakan merek kartu internet Indosat 9,38%, peluang responden akan tetap menggunakan merek kartu internet XL Axiata 0,72%, dan peluang responden akan tetap menggunakan merek kartu internet Hutchison 59,17%.

## V. KESIMPULAN DAN SARAN

### Kesimpulan

Berdasarkan analisa dan pembahasan yang telah dilakukan, kesimpulan dari penelitian ini adalah:

1. Peluang perpindahan merek kartu internet GSM Prabayar di kalangan mahasiswa S1 Matematika FMIPA UNIB saat ini dapat dilihat pada matriks peluang transisi  $P$  berikut:

$$P = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \end{matrix} & \begin{bmatrix} 0,5968 & 0,1129 & 0,0000 & 0,2903 \\ 0,3333 & 0,2500 & 0,0000 & 0,4167 \\ 0,0000 & 0,0000 & 0,0000 & 1,0000 \\ 0,1566 & 0,0602 & 0,0121 & 0,7711 \end{bmatrix} \end{matrix}$$

Keterangan:

- 0 = Responden Telkomsel
- 1 = Responden Indosat
- 2 = Responden XL Axiata
- 3 = Responden Hutchison

2. Prediksi peluang perpindahan pada masing-masing merek kartu internet GSM prabayar di kalangan mahasiswa S1 Matematika FMIPA UNIB, pada masa yang akan datang mengalami keadaan seimbang saat langkah ke 16 atau pada bulan Desember 2016. Artinya, peluang responden yang akan tetap menggunakan kartu internet Telkomsel 30,73%, kartu internet Indosat 9,38%, kartu internet XL Axiata 0,72%, dan kartu internet Hutchison 59,17%, nilai ini tidak akan berubah pada periode berikutnya.

### Saran

Penelitian selanjutnya dapat menggunakan faktor-faktor yang mempengaruhi perpindahan merek kartu internet GSM prabayar atau alasan konsumen melakukan perpindahan dari suatu merek ke merek yang lain.

## DAFTAR PUSTAKA

*Universitas Bina Nusantara, Jakarta. Vol.3,  
No.2 November 2012.*

Anonim. 2015. Telekomunikasi Seluler di Indonesia.  
[https://id.wikipedia.org/wiki/Telekomunikasi\\_seluler\\_di\\_Indonesia](https://id.wikipedia.org/wiki/Telekomunikasi_seluler_di_Indonesia). Tanggal Akses: 5 April 2015.

Anonim. 2015. Global System for Mobile Communication.  
[http://id.wikipedia.org/wiki/Global\\_System\\_for\\_Mobile\\_Communications](http://id.wikipedia.org/wiki/Global_System_for_Mobile_Communications). Tanggal Akses: 21 April 2015.

Hillier, F. S dan G. J. Lieberman. 2001. *Introduction to Operations Research (Seventh Edition)*. McGraw-Hill Higher Education, New York.

Sri, K dan Nuraeni. 2014. Faktor-faktor yang Mempengaruhi Perilaku *Brand Switching* pada Kartu Prabayar XL. *Jurnal Sketsa Bisnis*. Vol.1, No.1 Agustus 2014.

Langi, Y. A. R. 2001. Penentuan Klasifikasi State pada Rantai dengan Menggunakan Nilai Eigen dari Matriks Peluang Transisi. *Jurnal Ilmiah Sains*. Vol.11, No.1 April 2011.

Debora, N. R dan Yoestini. 2012. Analisis Pengaruh Ketidakpuasan Konsumen, Harga, dan Kebutuhan Mencari Variasi Terhadap Perpindahan Merek Sabun Lifebuoy Di Semarang. *Jurnal Manajemen Fakultas Ekonomi dan Bisnis Universitas Diponegoro*. Vol.1, No.2 2012.

Prastya, S. 2013. Analisis Faktor-faktor yang Mempengaruhi Perpindahan Merek (*Brand Switching*) Kartu Indosat IM3. *Skripsi*. Fakultas Ekonomika dan Bisnis Universitas Diponegoro Semarang.

Ribhan. 2006. Faktor-faktor yang Mempengaruhi *Brand Switching* pada Pengguna SIM Card di Fakultas Ekonomi Universitas Lampung. *Jurnal Bisnis dan Manajemen, Universitas Lampung, Bandar Lampung*. Vol.3, No.1 September 2006.

Ross, S.M. 2010. *Introduction to Probability Models (Tenth Edition)*. Elsevier Inc.

Lim, S dan H, Sarjono. 2012. Penentuan Pangsa Pasar Shampo Favorit dengan Bantuan Program QS for Windows. *Jurnal Ekonomi*

Winston, W. L. 2004. *Operations Research Application and Algorithm (Third Edition)*. Indiana University.

# ANALISIS REGRESI LOGISTIK ORDINAL FAKTOR-FAKTOR YANG MEMPENGARUHI KECELAKAAN LALU LINTAS TERHADAP TINGKAT KEPARAHAN KORBAN DI PROVINSI BENGKULU

Puce Angreni<sup>1</sup>, Sigit Nugroho<sup>2</sup>, dan Pepi Novianti<sup>3</sup>  
Program Studi Matematika  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Universitas Bengkulu  
Email: puce\_angreni@yahoo.com

## ABSTRAK

Kepadatan lalu lintas yang terjadi di Provinsi Bengkulu mengakibatkan tingginya angka kecelakaan pada kendaraan bermotor. Kecelakaan lalu lintas yang terjadi di Provinsi Bengkulu menimbulkan kerugian yang cukup besar. Hal ini terbukti dari tingginya angka kecelakaan lalu lintas yang tercatat di Kantor Kepolisian Daerah Provinsi Bengkulu setiap tahunnya. Berdasarkan Permasalahan tersebut, maka penelitian ini dilakukan untuk mengetahui model Regresi logistik ordinal dan menentukan faktor-faktor yang berpengaruh signifikan terhadap keparahan korban kecelakaan lalu lintas di Provinsi Bengkulu. Variabel respon yang diteliti mempunyai skala ordinal yang terdiri dari 3 kategori yaitu meninggal dunia, luka berat dan luka ringan. Variabel prediktor yang digunakan berupa data kategorik dan rasio yaitu jenis kelamin, pendidikan, usia, profesi, jenis kecelakaan, jenis kendaraan, SIM korban dan waktu kecelakaan. Untuk melihat hubungan antara variabel prediktor dengan variabel respon digunakan metode regresi logistik ordinal. Hasil analisis regresi logistik ordinal diperoleh dua model fungsi logit yaitu:

$$\begin{aligned} \text{logit}_0(X) &= -12.953 - 13.131X_{2,2} + 1.221X_3 + 1.380X_{4,2} \\ \text{logit}_1(X) &= -12,687 - 13.131X_{2,2} + 1.221X_3 + 1.380X_{4,2} \end{aligned}$$

Berdasarkan model regresi logistik ordinal yang terbentuk, variabel respon yang berpengaruh signifikan terhadap keparahan korban kecelakaan lalu lintas di provinsi Bengkulu yaitu variabel pendidikan korban kategori SMP ( $X_{2,2}$ ), usia korban ( $X_3$ ) dan profesi korban kategori karyawan swasta ( $X_{4,2}$ )

**Kata Kunci :** Regresi Logistik Ordinal, Keparahan Korban, Variabel Respon, Variabel Prediktor, Model Logit.

## I. PENDAHULUAN

Kecelakaan lalu lintas merupakan kejadian yang terjadi pada jalan umum yang dilintasi berbagai jenis kendaraan. Kecelakaan ini dapat mengakibatkan luka ringan, luka berat hingga meninggal dunia pada manusia. Menurut Organisasi Kesehatan Dunia (WHO) kecelakaan lalu lintas menelan korban jiwa sekitar 1,2 juta manusia setiap tahun. Kecelakaan lalu lintas dinilai menjadi pembunuh terbesar ketiga di Indonesia, setelah penyakit koroner dan TBC (Fitriah, Mashuri dan Irhamah, 2012).

Tingginya Kecelakaan lalu lintas pada kendaraan bermotor di jalan raya terutama di Provinsi Bengkulu menimbulkan kerugian materil yang cukup besar. Hal ini terbukti dari banyaknya jumlah korban kecelakaan lalu lintas setiap tahunnya di Provinsi Bengkulu. Berdasarkan data kantor kepolisian daerah Provinsi Bengkulu diperoleh jumlah korban kecelakaan lalu lintas sejak tahun 2010 sampai 2014 berjumlah 2.643 orang. Secara rinci, pada 2010 sebanyak 623 orang, 2011 sebanyak 231 orang, 2012 ada 412 orang, 2013 sebanyak 477 orang dan 2014 sebanyak 450 orang (BPS, 2015).

Terjadinya kecelakaan lalu lintas pada umumnya diakibatkan perilaku pengemudi yang melanggar peraturan perundang-undangan lalu lintas yang ada. Perilaku tersebut dapat berupa mengemudi kendaraan dengan kecepatan yang tinggi, tidak memiliki surat izin mengemudi, melanggar rambu-rambu lalu lintas serta marka jalan dan berbagai bentuk pelanggaran lainnya (Dharma dan Edison, 2010).

Banyak faktor yang dapat mempengaruhi tingginya angka kecelakaan lalu lintas. Adapun faktor penyebab kecelakaan antara lain adalah: manusia, sarana dan prasarana (kendaraan dan jalan), lingkungan (Wicaksono, Fathurchoman dan Riyanto, 2014).

Penggolongan kecelakaan lalu lintas berdasarkan tingkat keparahan korban dapat diklasifikasikan sebagai berikut:

- Korban meninggal adalah korban yang pasti meninggal akibat kecelakaan lalu lintas dalam jangka waktu paling lama 30 hari setelah kecelakaan.
- Korban luka berat adalah korban menderita cacat tetap atau harus dirawat 30 hari sejak kecelakaan.

- c. Korban luka ringan adalah korban yang tidak masuk dalam pengertian korban meninggal dan korban luka berat (Mahawati dan Prasetya, 2013).

Tingkat keparahan korban memiliki skala ordinal atau memiliki tingkatan pada setiap kategori. Untuk mendapatkan hubungan antara faktor-faktor penyebab kecelakaan dan tingkat keparahan korban dapat digunakan analisis regresi logistik ordinal. Regresi Logistik Ordinal merupakan salah satu metode statistika untuk menganalisis variabel respon yang mempunyai skala data ordinal dan terdiri dari tiga kategori atau lebih (Wulandari, Salamah dan Susilaningrum, 2009). Penelitian dengan menggunakan analisis regresi logistik ordinal telah beberapa kali dilakukan. (Indawati, Kusumaningrum dan Maena, 2010) telah melakukan penelitian aplikasi regresi logistik ordinal untuk pemodelan dan klasifikasi huruf mutu mata kuliah metode statistika. Penelitian regresi logistik ordinal juga dilakukan oleh (Raharjanti dan Widiharih, 2005) yaitu model logit kumulatif untuk respon ordinal. Selain itu (Fitriah, Mashuri dan Irhamah, 2012) melakukan penelitian mengenai faktor-faktor yang mempengaruhi keparahan korban kecelakaan lalu lintas di Kota Surabaya dengan pendekatan bagging regresi logistik ordinal.

Berdasarkan uraian di atas maka penulis tertarik untuk melakukan penelitian terhadap tingkat keparahan korban kecelakaan lalu lintas di Provinsi Bengkulu menggunakan regresi logistik ordinal dengan judul “Analisis Regresi Logistik Ordinal Faktor-Faktor yang Mempengaruhi Kecelakaan Lalu Lintas terhadap Tingkat Keparahannya Korban di Provinsi Bengkulu”.

## II. TINJAUAN PUSTAKA

### A. Regresi Logistik

Regresi logistik digunakan jika variabel respon bersifat kategorik (nominal atau ordinal) dengan variabel-variabel prediktor bersifat kontinu maupun kategorik. Apabila variabel respon terdiri dari dua kategori maka metode regresi logistik yang diterapkan adalah regresi logistik biner. Model regresi logistik adalah model regresi yang setiap peubah terikat atau responnya mensyaratkan berupa peubah kategorik, sedangkan menurut Hosmer dan Lemeshow (2000) metode regresi logistik adalah suatu metode analisis statistika yang mendeskripsikan hubungan antara peubah respon yang memiliki dua kategori atau lebih dengan satu atau lebih peubah penjelas berskala kategori atau interval.

### B. Regresi Logistik Ordinal

Regresi logistik ordinal adalah suatu metode analisis yang digunakan untuk mengetahui hubungan antara variabel respon dengan variabel prediktor, dimana variabel respon berskala ordinal yang terdiri atas tiga kategori atau lebih dan skala pengukurannya bersifat tingkatan (Akbar, Mukarromah dan Paramita, 2010).

Model yang dipakai untuk regresi logistik ordinal adalah model logit kumulatif. Model logit kumulatif

merupakan model yang diperoleh dengan cara membandingkan peluang kumulatif yaitu peluang kurang dari atau sama dengan kategori respon ke- $j$  pada  $p$  variabel prediktor yang dinyatakan dalam vektor  $\mathbf{X}$ ,  $P(Y \leq j|\mathbf{X})$  dengan peluang lebih besar dari kategori respon ke- $j$ ,  $P(Y > j|\mathbf{X})$  (Hosmer dan Lemeshow, 2000). Peluang kumulatif  $P(Y \leq j|\mathbf{X})$  didefinisikan sebagai berikut (Agresti, 2007):

$$P(Y \leq j|\mathbf{X}) = \frac{e^{(\theta_j + \sum_{k=1}^p \beta_k x_k)}}{1 + e^{(\theta_j + \sum_{k=1}^p \beta_k x_k)}}$$

Berdasarkan definisi model logit kumulatif, maka secara matematis model kumulatif model dapat ditulis sebagai berikut:

$$\text{logit } P(Y \leq j|\mathbf{X}) = \ln \left( \frac{P(Y \leq j|\mathbf{X})}{P(Y > j|\mathbf{X})} \right)$$

Jadi, diperoleh model logit kumulatif regresi logistik ordinal adalah:

$$\text{logit } P(Y \leq j|\mathbf{X}) = \theta_j + \sum_{k=1}^p \beta_k x_k$$

### C. Estimasi Parameter

Estimasi parameter pada regresi logistik ordinal dapat dilakukan dengan menggunakan *Maksimum Likelihood Estimation* (MLE) (Sari dan Kusri, 2009). MLE adalah sebuah metode statistik yang digunakan untuk menaksir parameter-parameter model regresi logistik dengan mengestimasi parameter  $\beta$  dengan memaksimalkan fungsi *Likelihood*. Sehingga agar lebih mudah terlebih dahulu dibentuk  $\ln$  dari fungsi *likelihood*, kemudian mendiferensialkan  $\ln$  dari fungsi *likelihood* tersebut terhadap masing-masing parameter. Fungsi *Likelihood* dapat dinyatakan sebagai berikut (Indawati, Kusumaningrum dan Maena, 2010):

$$l(\beta) = \prod_{i=1}^n [\pi_0(x_i)^{y_{0i}} \pi_1(x_i)^{y_{1i}} \pi_2(x_i)^{y_{2i}}]$$

### D. Pengujian Parameter

Pengujian parameter model regresi logistik ordinal dilakukan untuk memeriksa pengaruh variabel prediktor terhadap variabel respon. Uji parameter yang dilakukan ada dua yaitu uji serentak dengan statistik uji *Likelihood Ratio Test* ( $G$ ) dan uji individu dengan statistik uji *Wald* ( $W$ ).

### E. Uji Kesesuaian Model

Setelah dilakukan uji serentak dan individu, uji selanjutnya yaitu uji kesesuaian model dengan menggunakan uji *Deviance*. Prosedur pengujianya seperti berikut:

- a. Hipotesis Pengujian

$H_0: \hat{\pi}_i = y_i$  atau model logit yang diperoleh layak untuk digunakan.

$H_1: \hat{\pi}_i \neq y_i$  atau model logit yang diperoleh tidak layak untuk digunakan.

- b. Statistik Uji

$$D = -2 \sum_{i=1}^n y_i \ln \left( \frac{\hat{\pi}_i}{y_i} \right) + (1 - y_i) \ln \left( \frac{1 - \pi_i}{1 - y_i} \right)$$

Dengan  $\hat{\pi}_i = \hat{\pi}(x_i)$  adalah peluang observasi ke- $i$  pada kategori ke- $j$ . Derajat bebas uji ini  $(n - p)$  dengan  $n$  jumlah kovariat dan  $p$  jumlah variabel prediktor.

c. Kriteria Penolakan

Tolak  $H_0$  jika  $D_{hitung} > \chi^2_{(\alpha; (n-p))}$  atau  $\alpha \leq 0.1$

F. Interpretasi Parameter

Interpretasi parameter dari suatu model adalah inferensi dari pengambilan kesimpulan berdasarkan pada koefisien parameter. Interpretasi regresi logistik ordinal dapat dijelaskan dengan *odd rasio*. Nilai *odd rasio* yaitu nilai yang menunjukkan perbandingan tingkat kecenderungan dari dua kategori dalam satu variabel prediktor dengan salah satu kategorinya dijadikan pembanding atau kategori dasar.

*Odd rasio* dari dua kategori  $X$  yaitu:

$$\psi(x_2, x_1) = \frac{\frac{P[Y \leq j | x_2]}{P[Y > j | x_2]}}{\frac{P[Y \leq j | x_1]}{P[Y > j | x_1]}} = \frac{\exp(\beta_{0j} + \beta_1(x_2))}{\exp(\beta_{0j} + \beta_1(x_1))}$$

$$= e^{(\beta_{0j} + \beta_1(x_2))} \cdot \frac{1}{e^{(\beta_{0j} + \beta_1(x_1))}}$$

$$= e^{(\beta_{0j} + \beta_1(x_2)) - (\beta_{0j} + \beta_1(x_1))}$$

$$= e^{(\beta_{0j} + \beta_1(x_2) - \beta_{0j} - \beta_1(x_1))}$$

$$= e^{\beta_1(x_2) - \beta_1(x_1)}$$

$$= e^{\beta_1(x_2 - x_1)}$$

Log dari *odd rasio* pada persamaan (2.13) adalah

$$\ln e^{(\beta_1(x_2 - x_1))}$$

$$= \beta_1(x_2 - x_1) \quad (2.14)$$

Nilai *odd rasio* menginterpretasikan bahwa peluang respon pada kategori suatu dari atau sama dengan  $j$  dibandingkan dengan suatu respon pada kategori  $(j + 1)$  sampai dengan  $p$  untuk  $\mathbf{X} = x_2$  sebesar  $(\beta_1(x_2 - x_1))$  kali dari  $\mathbf{X} = x_1$  (Wulandari, Salamah dan Susilaningrum, 2009).

### III. METODOLOGI PENELITIAN

a. Sumber Data Penelitian

Data yang digunakan pada penelitian ini merupakan data sekunder yang diperoleh dari catatan kantor Kepolisian Daerah Provinsi Bengkulu. Variabel yang digunakan pada penelitian ini adalah data kecelakaan lalu lintas yang terjadi di Provinsi Bengkulu selama bulan Januari sampai September tahun 2015.

b. Variabel Penelitian

Variabel yang digunakan pada penelitian ini adalah variabel tentang keparahan korban kecelakaan lalu lintas di Provinsi Bengkulu sebagai variabel respon (Y) dan faktor-faktor penyebab kecelakaan lalu lintas di Provinsi Bengkulu sebagai variabel prediktor (X).

c. Teknik Analisis Data

Variabel yang digunakan pada penelitian ini adalah variabel tentang keparahan korban kecelakaan lalu lintas di Provinsi Bengkulu sebagai variabel respon (Y) dan faktor-faktor penyebab kecelakaan lalu lintas di Provinsi Bengkulu sebagai variabel prediktor (X).

### IV. HASIL DAN PEMBAHASAN

a. Deskripsi Data

Data yang digunakan adalah data kecelakaan lalu lintas yang tercatat di kantor Kepolisian Daerah Provinsi Bengkulu dari bulan Januari sampai dengan September 2015, yaitu sebanyak 189 data (Lampiran 1). Analisis ini menggunakan variabel keparahan korban kecelakaan lalu lintas di Provinsi Bengkulu sebagai variabel respon (Y) dan faktor-faktor penyebab kecelakaan lalu lintas sebagai variabel prediktor (X).

Variabel respon yang digunakan meliputi tingkat keparahan korban kecelakaan lalu lintas yang terdiri dari meninggal dunia, luka berat, dan luka ringan dengan frekuensi seperti Tabel 4.1 berikut:

Tabel 4.1 Distribusi Frekuensi Variabel Respon

Variabel	Frekuensi	Persentase	Persentase Valid	Persentase Kumulatif
Meninggal Dunia	16	8,5	8,5	8,5
Luka Berat	4	2,1	2,1	10,6
Luka Ringan	169	89,4	89,4	100,0
Total	189	100,0	100,0	

Berdasarkan Tabel.1 terlihat bahwa tingkat keparahan korban tertinggi yaitu luka ringan sebanyak 169 orang atau sebesar 89,4%. Tingkat keparahan korban meninggal dunia sebanyak 16 orang atau sebesar 8,5% dan tingkat keparahan luka berat sebanyak 4 orang atau sebesar 2,1% dari jumlah korban kecelakaan lalu lintas yang terjadi. Pada penelitian ini, faktor penyebab kecelakaan lalu lintas menggunakan variabel prediktor meliputi jenis kelamin, pendidikan korban, usia korban, profesi korban, jenis kecelakaan, waktu kecelakaan, jenis kendaraan dan SIM korban.

b. Model Terbaik

Berdasarkan hasil output SPSS untuk variabel yang telah signifikan diperoleh hasil estimasi parameter setelah pengujian ulang bahwa nilai  $p\_value$  masing-masing variabel kurang dari  $\alpha = 0,1$  maka dapat disimpulkan bahwa model yang diperoleh sudah baik. Sehingga persamaan model regresi logistik ordinal faktor-faktor yang mempengaruhi kecelakaan lalu lintas di Provinsi Bengkulu adalah sebagai berikut:

$$\text{logit}_0(X) = -12,953 - 13,131X_{2,2} + 1,221X_3 + 1,380X_{4,2}$$

$$\text{logit}_1(X) = -12,687 - 13,131X_{2,2} + 1,221X_3 + 1,380X_{4,2}$$

Keterangan:

$X_{2,2}$ : pendidikan korban (SMP)

$X_3$ : usia korban

$X_{4,2}$ : profesi korban (karyawan swasta)

c. Uji Kesesuaian Model

Uji kesesuaian model dilakukan untuk melihat apakah model regresi logistik ordinal yang diperoleh telah sesuai atau belum. Berikut adalah hasil uji kesesuaian model menggunakan uji *Devians*.

Tabel 4.7 Uji Kesesuaian Model

	Kai_Kuadrat	Db	Sig.
Pearson	13,107	42	1,000
Devians	11,012	42	1,000

Berdasarkan Tabel 4.7 dapat dilihat bahwa nilai uji *Devians*  $11,012 < \chi^2_{(0,1;42)} = 54,09$  dan nilai signifikansi  $1,00 > \alpha = 0,1$  sehingga  $H_0$  diterima, ini menunjukkan bahwa model yang telah dihasilkan layak untuk digunakan.

d. Interpretasi Parameter

Interpretasi parameter dapat dibentuk menggunakan *Odds Rasio*. Nilai *Odds Rasio* dari variabel yang signifikan dapat dilihat seperti tabel berikut ini:

Tabel 4.8 Nilai *Odds Rasio*

Variabel	Estimasi ( $\beta$ )	<i>Odds Rasio</i> ( $e^\beta$ )
( $X_{2,2}$ ) [Pendidikan_Korban=2]	-13,131	1,9828E-06
( $X_3$ ) Usia_Korban	1,221	3,39
( $X_{4,2}$ ) [Profesi_Korban=2]	1,380	3,97

V. KESIMPULAN DAN SARAN

a. Kesimpulan

Berdasarkan hasil dan pembahasan dalam penelitian “analisis regresi logistik ordinal faktor-faktor yang mempengaruhi kecelakaan lalu lintas terhadap tingkat keparahan korban di Provinsi Bengkulu” dapat disimpulkan bahwa:

a. Model regresi logistik ordinal yang diperoleh dapat menjelaskan hubungan tingkat keparahan korban kecelakaan lalu lintas terhadap faktor-faktor yang mempengaruhi kecelakaan, dengan model regresi logistik ordinal sebagai berikut:

$$\text{logit}_0(X) = -12,953 - 13,131X_{2,2} + 1,221X_3 + 1,380X_{4,2}$$

$$\text{logit}_1(X) = -12,687 - 13,131X_{2,2} + 1,221X_3 + 1,380X_{4,2}$$

b. Faktor-faktor yang berpengaruh signifikan terhadap keparahan korban kecelakaan lalu lintas di Provinsi Bengkulu dengan model regresi logistik ordinal adalah variabel pendidikan korban kategori SMP ( $X_{2,2}$ ), usia korban ( $X_3$ ) dan profesi korban kategori karyawan swasta ( $X_{4,2}$ ).

b. Saran

1. Penelitian selanjutnya dapat menggunakan metode lain sebagai pembandingan regresi logistik ordinal, sehingga dapat diketahui metode mana yang memiliki ketepatan klasifikasi lebih baik.
2. Ada baiknya dilakukan analisis regresi logistik ordinal yang lebih kompleks dengan melibatkan lebih banyak variabel prediktor agar mendapatkan nilai analisis yang lebih baik.

DAFTAR PUSTAKA

Agresti, A., 2007, *An Introduction To Categorical Data Analysis*, Jhon Wiley and Sons, Kanada.

Ahriyanti, R.Y., 2012, Analisis Statistik Data Kecelakaan Lalu Lintas Kota Bengkulu Tahun 2011, *Skripsi tidak diterbitkan*, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Bengkulu, Bengkulu.

Akbar, S.J., A. Mukarromah dan L. Paramita., 2010, Bagging Regresi Logistik Ordinal pada Status Gizi Balita, *Media Statistika*, vol. 3, No.2, Desember 2010. 103-116.

Ambarwati, A.N., H. Kuswanto dan I. Zain, 2012, Pendekatan Cart dan Regresi Logistik pada Pola Tingkat Keparahan Korban Kecelakaan Lalu Lintas di Surabaya, *Prosiding Seminar Nasional Penelitian, Pendidikan dan Penerapan Mipa*.

BPS, 2015, *Berita Resmi Statistik*, No. 04/01/17/Th.VIII.

Dharma, A. dan B. Edison., 2010, Identifikasi Kecelakaan Lalu Lintas, *Jurnal Fakultas Teknik Universitas Pasir Pangaraian*.

Fitriah, W.W., M. Mashuri dan Irhamah, 2012, Faktor-Faktor yang Mempengaruhi Keparahan Korban Kecelakaan Lalu Lintas di Kota Surabaya dengan Pendekatan Bagging Regresi Logistik Ordinal, *Jurnal Sains dan Seni ITS*, hlm: D-253.

Hosmer D.W and S. Lemeshow, 2000, *Applied Logistic Regression*. Jhon Wiley and Sons, Inc. New York.

Indawati., D. Kusumaningrum dan I. Maena, 2010, Aplikasi Regresi Logistik Ordinal Multilevel Untuk Pemodelan dan Klasifikasi Huruf Mutu Mata Kuliah Metode Statistika, *Forum Statistika dan Komputasi*, vol 15 No.2.

Mahawati, E., dan J. Prasetya, 2013, Analisis Penggunaan Handphone Saat Berkendara terhadap Potensial Kecelakaan Lalu Lintas Pada Remaja di Semarang, *Seminar Nasional*

*Teknologi Informasi dan Terapan 2013 (SEMANTIK 2013).*

Pamungkas, S.P., 2014, Mengenal Perilaku Pengendara Kendaraan Dalam Upaya Mencegah Terjadinya Kecelakaan di Jalan Raya, *Teknis volume 9, Nomor 1, 1 April 2014 : 13-18.*

Raharjanti R.P dan T. Widiharih, 2005, Model Logit Kumulatif untuk Respon Ordinal, *Jurnal Matematika Vol. 8, No.3.*

Sari, M.Y dan D.E. Kusri, 2009, Penggunaan Analisis Regresi Logistik Untuk Mencari Probabilitas Turnover Intention Beserta Faktor-Faktor yang Mempengaruhinya. *J-Statistika Vol 3 No 1.*

Wicaksono, D., R.A Fathurchoman, B. Riyanto, 2014, Analisis kecelakaan lalu Lintas, *Jurnal Karya Teknik Sipil vol. 3.(1), Halaman 203-213.*

Wulandari, S.P., M. Salamah dan D. Susilaningrum, 2009, *Diktat Pengajaran Analisis Data Kualitatif*, ITS, Surabaya.

## ANALISIS REGRESI TOBIT

(Studi Kasus: Faktor-Faktor Yang Berpengaruh Terhadap Biaya Pengeluaran Konsumsi Pakaian Dikalangan Mahasiswa Matematika Angkatan 2010-2013 FMIPA Unib)

Erwin Sinurat<sup>1)</sup>, Sigit Nugroho<sup>2)</sup>, Etis Sunandi<sup>2)</sup>

<sup>1)</sup> Alumni Jurusan Matematika FMIPA Universitas Bengkulu

<sup>2)</sup> Staf Pengajar Jurusan Matematika FMIPA Universitas Bengkulu

### ABSTRAK

Penelitian ini bertujuan untuk menetapkan model faktor-faktor yang berpengaruh terhadap biaya pengeluaran konsumsi pakaian dikalangan Mahasiswa Matematika Angkatan 2010-2013 FMIPA UNIB dengan menggunakan Model Regresi Tobit Standar. Struktur data tentang pengeluaran konsumsi pakaian merupakan data tersensor (*censored data*), dimana terdapat nilai nol pada sebagian observasinya, sedangkan untuk sebagian observasi yang lainnya mempunyai nilai tertentu yang bervariasi. Penggunaan metode analisis regresi linier klasik dengan metode *Ordinary Least Square* (OLS) pada data tersensor akan menimbulkan bias dalam data. Dari pemodelan tersebut ada beberapa variabel yang dinyatakan berpengaruh terhadap variabel Pengeluaran konsumsi pakaian ( $Y$ ) mahasiswa matematika FMIPA UNIB yaitu variabel Asal Daerah ( $X_2$ ), Jumlah Uang Bulanan ( $X_7$ ) dan Meningkatkan kepercayaan diri ( $X_{12}$ ). Dari model yang dihasilkan tersebut diperoleh nilai determinasi  $R^2$  yang cukup kecil yaitu 22%, hal tersebut dimungkinkan karena sampel yang digunakan terlalu kecil dan variabel yang digunakan terlalu banyak.

**Kata Kunci:** Model Regresi Tobit, Data Tersensor, Pengeluaran Konsumsi Pakaian

### I. PENDAHULUAN

Terdapat beberapa model regresi, namun pembedanya sendiri dilihat berdasarkan variabel terikat yang digunakan. Jika variabel terikat  $Y$  merupakan variabel kontinu, maka salah satu metode yang dapat digunakan adalah analisis regresi linier. Sedangkan pada data dengan variabel terikat berupa campuran atau *mixture* yaitu terdiri dari data kontinu dan pengukuran data diskrit maka digunakan regresi tobit.

Greene (2008) menyebutkan bahwa variabel terikat yang bersifat campuran (*mixture*) memiliki struktur data dengan skala diskrit untuk yang bernilai nol, dan berskala kontinu untuk yang tidak bernilai nol. Data tersebut disebut juga data tersensor. Sebaran data tersensor adalah sebaran normal tersensor, yang mengikuti asumsi  $N(\mu, \sigma^2)$ .

Penggunaan regresi tobit (regresi tersensor) pada data campuran akan mengurangi efek bias jika dibandingkan dengan data yang diolah menggunakan regresi linier klasik. Hal ini dikarenakan data yang bernilai nol (data diskrit) dapat diolah secara bersama dengan data kontinu sehingga tidak akan kehilangan informasi yang berasal dari data diskrit.

Pemodelan menggunakan data konsumsi juga sering ditemui permasalahan *zero expenditure* atau pengeluaran bernilai nol. Struktur data tentang pengeluaran konsumsi pakaian merupakan data tersensor (*censored data*), dimana terdapat nilai nol pada sebagian observasinya, sedangkan untuk sebagian observasi yang lainnya mempunyai nilai tertentu yang bervariasi.

### II. TINJAUAN PUSTAKA

#### A. Regresi Tobit

Regresi tobit pertama kali diperkenalkan oleh James Tobin pada tahun 1958. Regresi Tobit merupakan analisis regresi yang digunakan untuk variabel terikat yang sebagian datanya memiliki skala pengukuran diskrit dan sebagian yang lain berskala kontinu. Greene (2008) menyebutkan bahwa variabel terikat yang bersifat *mixture* (campuran) memiliki struktur data dengan skala diskrit untuk yang bernilai nol, dan berskala kontinu untuk yang tidak bernilai nol. Data tersebut disebut juga data tersensor. Tersensor sendiri dalam hal ini dapat berarti nilai dari variabel terikat tersebut terkonsentrasi atau terkelompok pada satu nilai (Novianti, 1993 dalam Salim, 2007). Sebaran data tersensor adalah sebaran normal tersensor, yang mengikuti asumsi  $N(\mu, \sigma^2)$ .

Model regresi tersensor merupakan salah satu metode statistika yang dapat digunakan untuk menentukan model bila terjadi pembatasan pada variabel terikatnya. Pada model regresi tersensor beberapa nilai sampel dicatat sebagai nilai batas dari nilai yang sebenarnya. Data pengamatan pada variabel jenis ini mengelompok akibat adanya batas bawah (tersensor kiri), batas atas (tersensor kanan) atau dapat juga keduanya. Pembatasan tersebut dapat terjadi secara alamiah seperti beberapa nilai yang lebih dekat terhadap suatu nilai tertentu. Pembatasan juga dapat ditentukan oleh peneliti tergantung pada tujuan penelitiannya (Frone, 1997).

Adanya pembatasan terhadap suatu nilai tertentu terhadap variabel terikat  $y$ , sebut saja  $a$ , mengakibatkan distribusi data tersebut berubah. Jika

suatu populasi telah diketahui berdistribusi normal, maka distribusi akibat adanya pemotongan nilai tertentu berubah menjadi distribusi normal tersensor, sehingga model menjadi:

$$y_i = \begin{cases} a, & \text{jika } y_i^* \leq a \\ y_i^*, & \text{jika } y_i^* > a \end{cases} \quad (1)$$

Sampel yang dihasilkan yaitu  $y_1, y_2, \dots, y_n$  disebut sampel tersensor.

Dalam model Tobit standar didefinisikan seperti masalah diatas dengan nilai  $a = 0$ . Formulasi model Tobit dalam Tobin (1958) dalam Greene (2008) secara umum adalah sebagai berikut.

$$y_i = \begin{cases} y_i^*, & \text{untuk } y_i^* > 0 \\ 0, & \text{untuk } y_i^* \leq 0 \end{cases} \quad (2)$$

Dimana  $i = 1, 2, \dots, T$  dan  $Y_i^*$  adalah variabel terikat dengan persamaan sebagai berikut.

$$y_i^* = \mathbf{x}_i^T \boldsymbol{\beta} + u_i \quad (3)$$

Dengan:

$y_i^*$  : adalah nilai variabel terikat yang sebenarnya  
 $\mathbf{x}_i^T$  :  $[1 \ x_{1i} \ x_{2i} \ \dots \ x_{pi}]$  adalah vektor variabel bebas

$\boldsymbol{\beta}$  :  $[\beta_0 \ \beta_1 \ \dots \ \beta_p]^T$  adalah vektor parameter,  $p$  merupakan banyaknya variabel

$u_i$  : residual model yang mengikuti sebaran normal tersensor  $(0, \sigma^2)$ .

## B. Distribusi Normal Tersensor

Variabel tersensor didefinisikan sebagai berikut, misalkan  $y^*$  berdistribusi normal dengan mean  $\mu$  dan varian  $\sigma^2$ :

$$y = \begin{cases} a, & \text{jika } y^* \leq a \\ y^*, & \text{jika } y^* > a \end{cases}$$

maka probabilitas tersensor  $y = a$  bernilai  $Prob(y = a) = Prob(y^* \leq a)$

$$= \int_{-\infty}^a \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{y^* - \mu}{\sigma}\right)^2\right) dy^*$$

$$= \int_{-\infty}^{\frac{a-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

$$\text{dengan } z = \frac{y^* - \mu}{\sigma} \text{ dan } \frac{1}{\sigma} dy^* = dz.$$

$$\text{Sehingga } Prob(y = a) = \Phi\left(\frac{a - \mu}{\sigma}\right).$$

Sedangkan untuk probabilitas tidak tersensor  $y = y^*$  adalah

$$\begin{aligned} Prob(y = y^*) &= Prob(y^* > a) \\ &= 1 - Prob(y^* \leq a) \\ &= 1 - \Phi\left(\frac{a - \mu}{\sigma}\right) \end{aligned}$$

Menurut Joreskog (2002) fungsi densitas dari  $y$  adalah:

$$\begin{aligned} f(y) &= \left[ \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} \right]^{1-j} \left[ \Phi\left(\frac{a-\mu}{\sigma}\right) \right]^j \\ &= \left[ \frac{1}{\sigma} \phi\left(\frac{a-\mu}{\sigma}\right) \right]^{1-j} \left[ \Phi\left(\frac{a-\mu}{\sigma}\right) \right]^j \end{aligned} \quad (4)$$

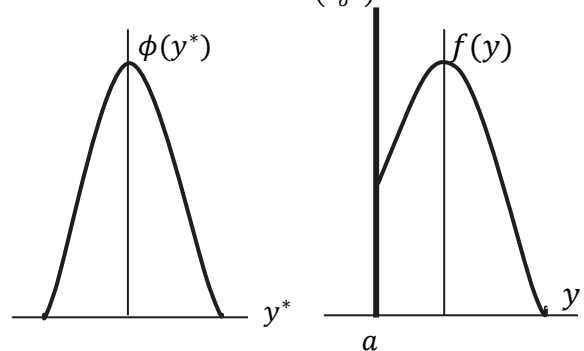
$$\text{dengan } j = \begin{cases} 1, & \text{jika } y = a \\ 0, & \text{jika lainnya} \end{cases}$$

$\phi$  dan  $\Phi$  masing-masing adalah fungsi densitas dan fungsi distribusi dari distribusi normal standar. Selanjutnya dari persamaan (4) diperoleh densitas untuk nilai  $y = y^*$  atau nilai  $y > a$  adalah:

$$f(y^*) = \frac{1}{\sigma} \phi\left(\frac{y^* - \mu}{\sigma}\right)$$

dan densitas dari  $y = a$  adalah :

$$f(a) = Prob(y = a) = \Phi\left(\frac{a-\mu}{\sigma}\right)$$



**Gambar 1.** Variabel Normal  $y^*$  dan Variabel Tersensor  $y$

Pada Gambar di sebelah kiri menunjukkan bahwa data menyebar mengikuti distribusi normal, sedangkan pada gambar disebelah kanan, akibat adanya batas terhadap suatu nilai maka data menyebar mengikuti sebaran distribusi normal tersensor namun dengan total peluang tetap bernilai satu.

## C. Fungsi Likelihood Model Tobit Standar

Fungsi Likelihood dari Model Tobit standar adalah:

$$L = \prod_{y_i=0}^{n_0} \left[ 1 - \Phi\left(\frac{x_i' \boldsymbol{\beta}}{\sigma}\right) \right] \prod_{y_i>0}^{n_1} \sigma^{-1} \phi[(y_i - x_i' \boldsymbol{\beta})/\sigma] \quad (5)$$

dimana:

$$\prod_{y_i=0}^{n_0} = \text{Adalah perkalian dari banyaknya pengamatan dimana } y_i = 0 \text{ atau } y_i^* \leq 0$$

$$\prod_{y_i>0}^{n_1} = \text{Adalah perkalian dari banyaknya pengamatan dimana } y_i > 0 \text{ atau } y_i^* > 0$$

misalkan:

$n_0$  = adalah banyaknya pengamatan dimana  $y_i = 0$

$n_1$  = adalah banyaknya pengamatan dimana  $y_i > 0$

dengan  $n_0 + n_1 = n$ ,  $\phi(\cdot)$  dan  $\Phi(\cdot)$  masing-masing menyatakan fungsi probabilitas densitas normal standar dan fungsi distribusi dari normal standar.

Sehingga,

1. Untuk  $y_i = 0$  atau  $y_i^* \leq 0$  didapatkan:

$$\Pr(Y = 0) = \Pr(Y^* \leq 0)$$

$$\Pr(Y = 0) = 1 - \Phi(x'_i\beta/\sigma) \quad (6)$$

Bila pemisalan yang dilakukan adalah  $u = \sigma t$ , dengan  $\sigma > 0$ , maka akan didapatkan  $\Pr(Y = 0) = \Phi(x'_i\beta/\sigma)$

2. Untuk  $y_i > 0$  atau  $y_i^* > 0$  diperoleh:

$$\Pr(Y, y_i > 0) = f(y_i^* | y_i^* > 0) \Pr(Y^* > 0)$$

$$= \frac{f(y_i^*)}{\Pr(Y^* > 0)} \Pr(Y^* > 0)$$

$$= f(y_i^*)$$

Karena  $Y^* \sim N(x'_i\beta, \sigma^2)$ , dan juga berlaku  $y_i = y_i^*$  maka

$$\Pr(Y, y_i > 0) = \sigma^{-1} \phi((y_i - x'_i\beta)/\sigma) \quad (7)$$

Fungsi Likelihood untuk model Tobit Standar adalah (persamaan (6))

$$L = \prod_{y_i=0}^{n_0} \Pr(Y, y_i = 0) \prod_{y_i>0}^{n_1} \Pr(Y, y_i > 0)$$

$$= \prod_{y_i=0}^{n_0} \left[ 1 - \Phi\left(\frac{x'_i\beta}{\sigma}\right) \right] \prod_{y_i>0}^{n_1} \sigma^{-1} \phi\left[\frac{y_i - x'_i\beta}{\sigma}\right]$$

dari persamaan (6) dan (7)

#### D. Estimasi Parameter Model Tobit

Dalam menduga parameter yang ada pada model tobit standar, yakni  $\beta$  dan  $\sigma^2$  berdasarkan  $n$  pengamatan  $y_i$ , tidak digunakan metode *Least Squares* karena kelinieran hubungan antara variabel terikat dengan variabel bebas tidak terpenuhi. Penggunaan metode Penaksiran *Least Squares* dalam menaksir parameter model tobit adalah bias. Oleh karena itu dalam menaksir parameter yang ada pada model tobit standar digunakan *Maksimum Likelihood*. Dengan mensubstitusikan persamaan (6) dan (7) maka persamaan Likelihood dari persamaan (5) dapat ditulis sebagai:

$$L = \prod_{y_i=0}^{n_0} \left[ 1 - \Phi\left(\frac{x'_i\beta}{\sigma}\right) \right] \prod_{y_i>0}^{n_1} \left[ \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(y_i - x'_i\beta)^2} \right] \quad (8)$$

Maka fungsi ln Likelihoodnya:

$$\ln L = \sum_{y_i=0}^{n_0} \ln \left[ 1 - \Phi\left(\frac{x'_i\beta}{\sigma}\right) \right] + \sum_{y_i>0}^{n_1} \ln \left[ \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(y_i - x'_i\beta)^2} \right] \quad (9)$$

Dimana:

$\sum_{y_i=0}^{n_0}$  = adalah penjumlahan dari banyaknya pengamatan dimana  $y_i = 0$

$\sum_{y_i>0}^{n_1}$  = adalah penjumlahan dari banyaknya pengamatan dimana  $y_i > 0$

misalkan:

$n_0$  = adalah banyaknya pengamatan dimana  $y_i = 0$

$n_1$  = adalah banyaknya pengamatan dimana  $y_i > 0$

dengan  $n_0 + n_1 = n$

Perhatikan suku terakhir dari persamaan (9)

$$= \sum_{y_i>0}^{n_1} \ln \left[ \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(y_i - x'_i\beta)^2} \right]$$

$$= -n_1 \ln \sigma - \frac{n_1}{2} \ln(2\pi) - \frac{1}{2\sigma^2} \sum_{y_i>0}^{n_1} (y_i - x'_i\beta)^2$$

$$= -\frac{n_1}{2} \ln \sigma^2 - \frac{n_1}{2} \ln(2\pi) - \frac{1}{2\sigma^2} \sum_{y_i>0}^{n_1} (y_i - x'_i\beta)^2 \quad (10)$$

Dengan mensubstitusikan persamaan (10) ke (9) diperoleh:

$$\ln L = \sum_{y_i>0}^{n_1} \ln \left[ 1 - \Phi\left(\frac{x'_i\beta}{\sigma}\right) \right] - \frac{n_1}{2} \ln \sigma^2$$

$$- \frac{n_1}{2} \ln(2\pi)$$

$$- \frac{1}{2\sigma^2} \sum_{y_i>0}^{n_1} (y_i - x'_i\beta)^2$$

Untuk mencari nilai ekstrimnya:

$$\frac{\partial \ln L}{\partial \beta} = 0 \text{ dan } \frac{\partial \ln L}{\partial \sigma^2} = 0, \text{ maka}$$

$$0 = -\frac{1}{\sigma} \sum_{y_i=0}^{n_0} \frac{\phi(x'_i\beta/\sigma)x'_i}{1 - \Phi(x'_i\beta/\sigma)} + \frac{1}{\sigma^2} \sum_{y_i>0}^{n_1} (y_i - x'_i\beta)x'_i \quad (11)$$

$$0 = -\frac{1}{2\sigma^3} \sum_{y_i=0}^{n_0} \frac{(x'_i\beta)\phi(x'_i\beta/\sigma)}{1 - \Phi(x'_i\beta/\sigma)} - \frac{n_1}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{y_i>0}^{n_1} (y_i - x'_i\beta)^2 \quad (12)$$

Oleh Olsen (1978) didapatkan matriks Hessiannya sebagai berikut:

$$= \begin{bmatrix} \sum_{y_i=0}^{n_0} \frac{\phi_i}{1-\Phi_i} \left( x_i' \alpha - \frac{\phi_i}{1-\Phi_i} \right) x_i' x_i & 0 \\ 0 & -\frac{n_1}{h^2} \end{bmatrix} - \begin{bmatrix} \sum_{y_i>0}^{n_1} x_i' x_i & -\sum_{y_i>0}^{n_1} x_i' y_i \\ -\sum_{y_i>0}^{n_1} x_i y_i & \sum_{y_i>0}^{n_1} y_i^2 \end{bmatrix} \quad (13)$$

Menyelesaikan persamaan (11) dan (12) maka akan diperoleh penaksir  $\beta$  dan  $\sigma^2$ . Untuk mendapatkan penaksir dari  $\sigma^2$ , kalikan persamaan (11) dengan  $\beta/(2\sigma^2)$  dan tambahkan dengan persamaan (12) akan didapatkan:

$$\sigma^2 = \frac{\sum_{y_i>0}^{n_1} (y_i - x_i' \beta) y_i}{n_1}$$

$$\beta = \beta_{LS} - \sigma (X_1' X_1)^{-1} X_1' y_0 \quad (14)$$

Persamaan (14) secara eksplisit memperlihatkan hubungan antara Penaksir Maksimum Likelihood untuk  $\beta$  Model Tobit dan Penaksir *Least Squares* yang diperoleh dari pengamatan tak nol  $y$ .

## E. Pengujian Parameter

### 1. Uji Serentak Menggunakan Uji G

Uji serentak digunakan untuk menguji parameter hasil dugaan secara bersama-sama. Pengujian menggunakan metode *likelihood ratio* atau uji  $G$ .

$$L(\omega) = \prod_{i=1}^n f(y_i; \beta_0) \text{ dengan } \omega = \{\beta_0\}$$

$$L(\Omega) = \prod_{i=1}^n f(y_i; (\beta_1, \beta_2, \dots, \beta_k))$$

dengan  $\Omega = \{(\beta_1, \beta_2, \dots, \beta_k)\}$

Hipotesis :

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

$$H_1 : \text{Paling tidak terdapat satu } \beta_j \neq 0, j = 1, 2, \dots, k$$

Statistik uji :

$$G = -2 \ln \left( \frac{L(\hat{\omega})}{L(\hat{\Omega})} \right) \quad (15)$$

Dimana :

$L(\hat{\omega})$  : Likelihood tanpa variabel bebas tertentu

$L(\hat{\Omega})$  : Likelihood dengan variabel bebas tertentu

$H_0$  ditolak jika uji  $G > \chi_{\alpha, k}^2$  berarti paling tidak satu  $\beta_j$  yang mempunyai peran terhadap model.

### 2. Uji Individu Menggunakan Uji Wald

Uji ini dilakukan untuk menguji setiap  $\beta_j$  secara individual untuk menunjukkan apakah suatu variabel bebas layak untuk masuk dalam model.

Hipotesis:

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0, \text{ untuk } j = 1, 2, \dots, 7$$

Statistik uji:

$$W^2 = \frac{\hat{\beta}_j^2}{(Se \hat{\beta}_j)^2} \quad (16)$$

Dimana:

$(Se \hat{\beta}_j)^2$  : standar error koefisien parameter

$\hat{\beta}_j^2$  : nilai koefisien dugaan variabel bebas mengikuti distribusi  $\chi^2$   $H_0$  ditolak jika  $W^2 > \chi_{\alpha, v}^2$  berarti  $\beta_j$  mempunyai peran terhadap model.

## F. Penentuan Kesesuaian Model

Menentukan kesesuaian model dalam analisis regresi salah satunya dapat menggunakan koefisien determinasi  $R^2$ . Penentuan kesesuaian model didasarkan pada besarnya nilai  $R^2$ . Pada kasus regresi tobit penulisan  $R^2$  dituliskan sebagai berikut (Bierens, 2004).

$$R^2 = 1 - \frac{\sum_{i=1}^n \hat{u}_i^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (17)$$

## G. Pengujian Sifat MLE Tobit dibawah Asumsi Ketidakstandaran

Setelah dilakukan pembentukan model regresi Tobit, selanjutnya akan dilakukan pengujian sifat-sifat dasar Metode MLE untuk model Tobit standar dibawah asumsi Ketidakstandaran (Amemiya, 1985).

Pengujiannya sama dengan pengujian pada uji asumsi klasik hanya saja hasilnya akan bertentangan dengan asumsi klasik pada regresi yaitu ketidaknormalan data, bersifat heteroskedastisitas dan mempunyai autokorelasi.

### 1. Uji Kenormalan data

Pengujian normalitas digunakan untuk melihat residual atau galat terdistribusi normal atau tidak. Untuk menguji asumsi kenormalan galat digunakan uji *Kolmogorov Smirnov*. Uji *Kolmogorov-Smirnov* bekerja dengan cara membandingkan dua buah distribusi/sebaran data, yaitu distribusi yang dihipotesiskan dan distribusi yang teramati.

Hipotesis dalam uji normalitas adalah:

$$H_0 : \text{Data menyebar normal}$$

$$H_1 : \text{Data tidak menyebar normal}$$

Statistik uji yang digunakan adalah:

$$D = \max |F_0(x) - S_N(x)| \quad (18)$$

$H_0$  ditolak jika nilai  $D$  hitung lebih besar dari nilai  $D$  tabel.

### 2. Uji Multikolinieritas

Multikolinieritas adalah adanya hubungan linier yang sempurna diantara beberapa atau semua peubah bebas dari model regresi berganda (Gujarati, 2004). Pendeteksian multikolinieritas dilakukan dengan melihat nilai *variance inflation factor* (VIF) dan *tolerance* (TOL) dengan ketentuan jika nilai VIF melebihi angka 10, maka terjadi multikolinieritas dalam model regresi. Kemudian jika nilai TOL sama

dengan 1, tidak terjadi multikolinieritas dalam model regresi (Kutner *et al.*, 2004).

### 3. Uji Heteroskedastisitas

Heteroskedastisitas adalah variansi dari error model regresi tidak konstan atau variansi antar error yang satu dengan error yang lain berbeda (Widarjono, 2007). Salah satu cara yang dapat digunakan mendeteksi adanya heteroskedastisitas dengan Metode Glejser. Glejser mengatakan bahwa nilai variansi variabel error model regresi tergantung dari variabel bebas. Selanjutnya untuk mengetahui apakah pola variabel error mengandung heteroskedastisitas Glejser menyarankan untuk melakukan regresi nilai mutlak residual dengan variabel bebas. Jika hasil uji F dari model regresi yang diperoleh tidak signifikan, maka tidak ada heteroskedastisitas dalam model regresi (Widarjono, 2007).

### 4. Uji Autokorelasi

Adanya autokorelasi pada error mengindikasikan bahwa ada satu atau beberapa faktor (variabel) penting yang mempengaruhi variabel terikat  $Y$  yang tidak dimasukkan ke dalam model regresi. Statistik uji yang sering dipakai adalah Durbin-Watson statistics (DW-statistics).

Kemudian Durbin-Watson berhasil menurunkan nilai kritis batas bawah ( $d_L$ ) dan batas atas ( $d_U$ ) sehingga jika nilai  $d$  hitung terletak di luar nilai kritis ini, maka ada atau tidaknya autokorelasi baik positif atau negatif dapat diketahui.

## H. Uji Validitas dan Reabilitas

### 1. Uji Validitas.

Pengujian validitas digunakan untuk mengukur sah/valid atau tidaknya butir kuesioner. Kuesioner dikatakan valid jika butir pertanyaan kuesioner mampu untuk mengungkapkan sesuatu yang akan diukur. Mencari korelasi antara skor butir pertanyaan dan total skor konstruk (Arikunto, 2006). Dalam menentukan signifikan atau tidak signifikan dengan membandingkan nilai  $r$  hitung dengan nilai  $r$  tabel *degree of freedom* =  $n - k$ , dan daerah sisi pengujian dengan alpha 0,05. Jika  $r$  hitung tiap butir pertanyaan bernilai positif dan lebih besar terhadap  $r$  tabel maka butir pertanyaan tersebut dikatakan valid.

### 2. Uji Reliabilitas

Reliabilitas adalah untuk mengukur suatu kuesioner yang merupakan indikator dari variabel atau konstruk. Butir pertanyaan dikatakan reliabel atau andal apabila jawaban seseorang terhadap pertanyaan adalah konsisten. Arikunto (2006) menyatakan "Reliabilitas menunjuk pada suatu pengertian bahwa sesuatu instrumen cukup dapat dipercaya untuk digunakan sebagai alat pengumpul data karena instrumen tersebut sudah baik". Apabila koefisien *Cronbach Alpha* ( $\alpha$ )  $\geq 0.7$  maka dapat dikatakan instrumen tersebut reliabel.

## III. METODE PENELITIAN

### A. Jenis Penelitian

Jenis penelitian yang digunakan dalam penulisan skripsi ini adalah studi kasus atau penelitian terapan (*applied research*) yang memiliki tujuan untuk menerapkan metode analisis yang sesuai dengan tujuan masing-masing analisis dalam suatu studi kasus yang mendukung.

### B. Populasi dan Sampel

Adapun yang menjadi populasi dalam penelitian ini adalah Mahasiswa Matematika FMIPA Angkatan 2010-2013 Universitas Bengkulu (Unib). Penelitian menggunakan teknik *Stratified Random Sampling*, yaitu cara mengambil sampel dengan memperhatikan strata (tingkatan) di dalam populasi. Pada penelitian ini sampel diambil dari mahasiswa Matematika FMIPA Universitas Bengkulu angkatan 2010-2013. Mahasiswa angkatan 2010 berjumlah 61, angkatan 2011 berjumlah 32, angkatan 2012 berjumlah 36 dan angkatan 2013 berjumlah 43. Jadi total seluruh mahasiswa angkatan 2010 hingga 2013 adalah 172 mahasiswa.

### C. Variabel Penelitian

#### 1. Variabel Terikat

Dalam penelitian ini digunakan biaya pengeluaran konsumsi pakaian dikalangan mahasiswa sebagai variabel terikat. Biaya pengeluaran konsumsi pakaian menunjukkan harga yang dikeluarkan untuk membeli pakaian, dalam selang waktu satu bulan terakhir. Pengukuran data berupa (1) skala diskrit, Rp.0 untuk menyatakan tidak ada biaya yang dikeluarkan dan (2) skala kontinu menyatakan besarnya biaya pengeluaran responden tergantung pada seberapa besar biaya pengeluaran terhadap konsumsi pakaian.

#### 2. Variabel Bebas

Variabel bebas merupakan yang besarnya tidak tergantung pada variabel lain. Berikut variabel-variabel bebasnya.

Jenis Kelamin ( $X_1$ ), Asal Daerah ( $X_2$ ), Uang Sumber lain selain dari orang tua ( $X_3$ ), Penghasilan Orang tua ( $X_4$ ), Tempat Tinggal ( $X_5$ ), Jumlah Saudara ( $X_6$ ), Jumlah Uang Bulanan ( $X_7$ ), Pekerjaan Ayah ( $X_8$ ), Status Sosial ( $X_9$ ), Gaya hidup Sosialita ( $X_{10}$ ), Komunikasi peran status ( $X_{11}$ ), meningkatkan kepercayaan diri ( $X_{12}$ ), teman, keluarga berperan ( $X_{13}$ ), Berapa Sering Online ( $X_{14}$ ), Lingkungan Berpengaruh ( $X_{15}$ ), Lebih suka online daripada toko ( $X_{16}$ ), Tertarik dengan diskon ( $X_{17}$ ), Lebih suka pakaian bermerk ( $X_{18}$ ), Lebih suka pakaian dipromosikan ( $X_{19}$ ) dan Harga merupakan keputusan akhir ( $X_{20}$ ).

#### 3. Alat Pengumpulan Data Primer

Dalam penelitian ini alat pengumpulan data yang digunakan adalah Kuesioner Digital, dengan

jumlah yang ditanyakan kepada responden sebanyak 20 pertanyaan.

#### IV. HASIL DAN PEMBAHASAN

Pada uji validitas dan reliabilitas, jumlah sampel yang digunakan adalah sebanyak 30 responden dari sampel keseluruhan 120 responden atau sebesar 25% dari total keseluruhan. Sedangkan karakteristik responden terdiri dari 23 orang dengan jenis kelamin perempuan dan 7 laki-laki. Hal ini dimaksudkan agar dengan jumlah minimal ini, distribusi skor/nilai akan lebih mendekati kurva normal (Arikunto, 2006). Data ini dikumpulkan menggunakan kuesioner digital yang disebar melalui jejaring sosial media dan calon responden dikontrol hanya boleh mengisi kuesioner sebanyak satu kali.

##### A. Uji Validitas dan Reliabilitas

Dari 13 variabel yang diuji, disimpulkan bahwa hanya ada 4 (empat) variabel yang dinyatakan dipertahankan atau dilanjutkan dalam pengujian selanjutnya karena asumsi validitas dan reliabilitas terpenuhi. Sedangkan untuk variabel lain diharapkan dibuang atau diperbaiki guna validitas dan reliabilitasnya terpenuhi secara bersama-sama. Uji validitas dilihat berdasarkan nilai korelasi Spearman yang dibandingkan dengan  $r$ -tabel, sedangkan reliabilitas dilihat berdasarkan nilai cronbach alpa yang lebih dari 0.7 maka dikatakan reliabel. Selain dari keempat variabel yang sudah dinyatakan valid dan reliabel variabel lain yang akan diuji adalah delapan variabel pokok, sehingga total variabel yang akan diuji pada tahap selanjutnya adalah 12 variabel.

**Tabel 1.** Variabel yang dinyatakan valid dan reliabel

Variabel	Validitas	Reliabilitas	Keputusan
Tempat Tinggal ( $X_5$ )	Valid	Tidak Reliabel	Dieliminasi
Status Sosial ( $X_9$ )	Valid	Reliabel	Dipertahankan
Gaya hidup Sosialita ( $X_{10}$ )	Valid	Reliabel	Dipertahankan
Komunikasi peran status ( $X_{11}$ )	Valid	Reliabel	Dipertahankan
Meningkatkan kepercayaan diri ( $X_{12}$ )	Valid	Reliabel	Dipertahankan
Temannya, keluarga berperan ( $X_{13}$ )	Valid	Tidak Reliabel	Dieliminasi
Berapa Sering Online ( $X_{14}$ )	Valid	Tidak Reliabel	Dieliminasi
Lingkungan Berpengaruh ( $X_{15}$ )	Tidak Valid	Tidak Reliabel	Dieliminasi
Online/Toko ( $X_{16}$ )	Tidak Valid	Tidak Reliabel	Dieliminasi
Tertarik dengan diskon ( $X_{17}$ )	Valid	Tidak Reliabel	Dieliminasi
Pakaian bermerk ( $X_{18}$ )	Valid	Tidak Reliabel	Dieliminasi
Pakaian dipromosikan ( $X_{19}$ )	Valid	Tidak Reliabel	Dieliminasi
Harga keputusan akhir ( $X_{20}$ )	Tidak Valid	Tidak Reliabel	Dieliminasi

Dari 13 variabel yang diuji, berdasarkan Tabel 1 disimpulkan bahwa hanya ada 4 (empat) variabel yang dinyatakan dipertahankan atau dilanjutkan dalam pengujian selanjutnya karena asumsi validitas dan reliabilitas terpenuhi. Sedangkan untuk variabel lain diharapkan dibuang atau diperbaiki guna validitas dan reliabilitasnya terpenuhi secara bersama-sama. Selain dari keempat variabel yang sudah dinyatakan valid dan reliabel variabel lain yang akan diuji adalah delapan variabel pokok, sehingga total variabel yang akan diuji pada tahap selanjutnya

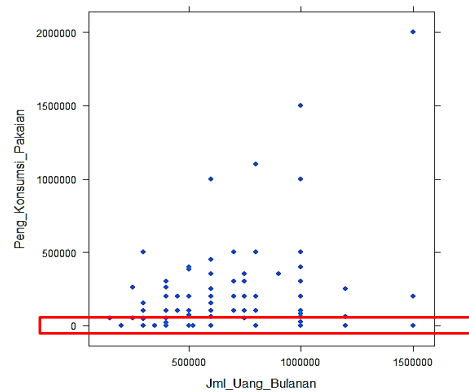
adalah 12 variabel. Adapun variabel yang akan diuji pada tahap selanjutnya diringkas pada Tabel 2.

**Tabel 2.** Variabel yang akan diuji tahap selanjutnya

No	Variabel	Jenis Data
1.	Jenis Kelamin ( $X_1$ )	Nominal
2.	Asal Daerah ( $X_2$ )	Nominal
3.	Uang Sumber lain selain dari tua ( $X_3$ )	Interval
4.	Penghasilan orang tua ( $X_4$ )	Ordinal
5.	Tempat Tinggal ( $X_5$ )	Nominal
6.	Jumlah Saudara ( $X_6$ )	Interval
7.	Jumlah Uang Bulanan ( $X_7$ )	Interval
8.	Pekerjaan Ayah ( $X_8$ )	Nominal
9.	Status Sosial ( $X_9$ )	Ordinal
10.	Gaya hidup Sosialita ( $X_{10}$ )	Ordinal
11.	Komunikasi peran status ( $X_{11}$ )	Ordinal
12.	Meningkatkan kepercayaan diri ( $X_{12}$ )	Ordinal

##### B. Identifikasi Variabel

Seperti pada penjelasan sebelumnya, bahwa regresi tobit merupakan analisis regresi yang digunakan untuk variabel terikat yang sebagian datanya memiliki skala pengukuran diskrit dan sebagian yang lain berskala kontinu yang disebut juga data tersensor. Tersensor sendiri dalam hal ini dapat berarti nilai dari variabel terikat tersebut terkonsentrasi atau terkelompok pada satu nilai, dalam kasus ini pula data terkonsentrasi pada pengamatan bernilai nol yang tampak seperti Gambar 1 berikut.



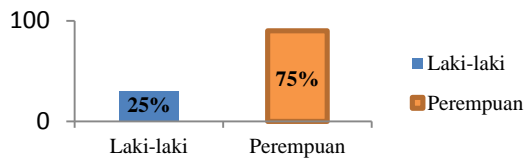
**Gambar 2.** Variabel Y tersensor di titik nol

Gambar di atas menunjukkan hubungan antara salah satu variabel bebas yaitu jumlah uang bulanan ( $X_7$ ) dengan variabel terikatnya yaitu pengeluaran konsumsi pakaian. Berdasarkan Gambar 2 dapat dilihat terdapat observasi yang terkonsentrasi pada nol yaitu untuk variabel pengeluaran konsumsi pakaian ( $Y$ ). Berarti dalam hal ini ada observasi variabel terikat yang terkonsentrasi pada suatu nilai, yaitu titik 0 (nol).

##### C. Deskripsi Data

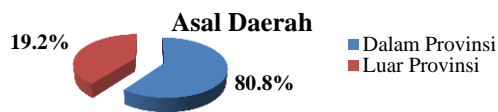
Pembahasan mengenai deskripsi data pada penelitian ini meliputi jumlah jawaban responden pada setiap pertanyaan dan persentase jawaban

responden. Uraian dari masing-masing variabel akan dijelaskan pada grafik dan diagram di bawah.



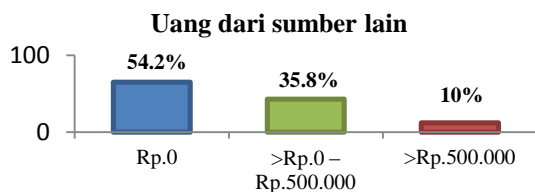
**Gambar 3.** Grafik Distribusi variabel Jenis Kelamin

Pada Grafik di atas menunjukkan bahwa jenis kelamin responden penelitian yang lebih dominan adalah perempuan sebanyak 90 orang atau 75% sedangkan jenis kelamin laki-laki sebanyak 30 orang atau dengan persentase 25% dari total keseluruhan.



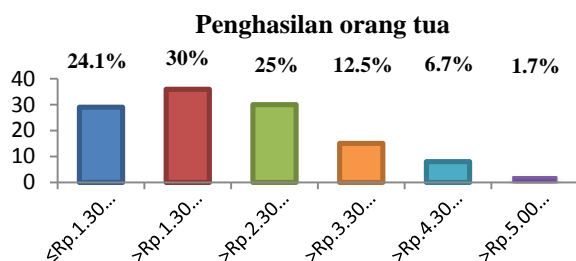
**Gambar 4.** Diagram Distribusi variabel Asal Daerah

Pada Gambar 4 menunjukkan bahwa responden didominasi oleh mahasiswa yang berasal dari dalam Provinsi Bengkulu yaitu sebanyak 97 orang atau 80.8% dari total keseluruhan sedangkan mahasiswa yang berasal dari luar Provinsi Bengkulu sebanyak 23 orang atau 19.2%.



**Gambar 5.** Distribusi variabel Uang dari sumber lain

Pada Gambar 4 menunjukkan bahwa variabel uang dari sumber lain selain dari orang tua responden penelitian dominannya adalah mahasiswa yang tidak mempunyai biaya pemasukan lain selain dari orang tua atau uang dari sumber lainnya nol rupiah yaitu sebanyak 65 orang atau 54.2% sedangkan yang mempunyai biaya pemasukan lain selain dari orang tua dapat dilihat pada Gambar 5.

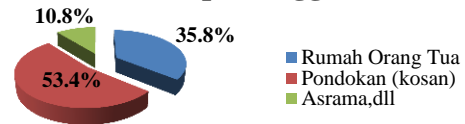


**Gambar 6.** Distribusi variabel Penghasilan Orang tua

Pada Gambar 6 menunjukkan bahwa penghasilan orang tua responden penelitian yang paling dominan adalah penghasilan sebesar >Rp.1.300.000– Rp.2.300.000 yaitu sebanyak 36

orang atau 30% sedangkan penghasilan dengan kategori penghasilan orang tua paling sedikit adalah yang penghasilan orang tuanya >Rp.5.000.000 yaitu sebanyak 2 orang atau 1.7% dari total keseluruhan.

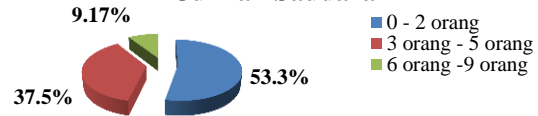
### Tempat tinggal



**Gambar 7.** Diagram variabel Tempat tinggal

Pada Gambar 7 menunjukkan bahwa tempat tinggal responden penelitian selama kuliah yang dominan adalah mahasiswa yang tinggal dipondokkan (kosan) yaitu sebanyak 64 orang atau 53.4% dari total keseluruhan.

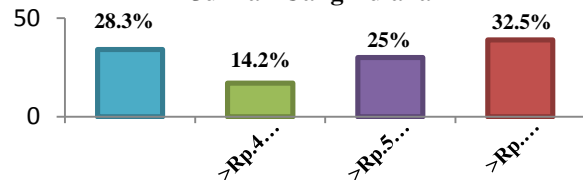
### Jumlah Saudara



**Gambar 8.** Diagram variabel Jumlah Saudara

Pada Gambar 8 menunjukkan bahwa variabel jumlah saudara didominasi oleh responden penelitian yang mempunyai jumlah saudara antara 0 - 2 orang yaitu sebanyak 64 orang atau 53.3% dari total keseluruhan, sedangkan responden yang mempunyai jumlah saudara dengan kategori lain dapat dilihat pada Gambar 8.

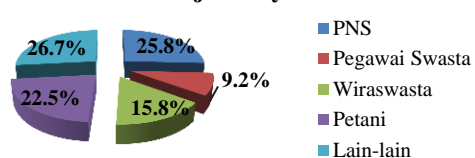
### Jumlah Uang Bulanan



**Gambar 9.** Distribusi variabel Jumlah uang bulanan

Pada Gambar 8 menunjukkan bahwa variabel jumlah uang bulanan didominasi oleh responden penelitian yang mempunyai jumlah uang bulanan sebesar >Rp.750.000 yaitu sebanyak 39 orang atau 32.5% dari total keseluruhan, sedangkan responden yang mempunyai uang bulanan dengan kategori lain dengan jelas ditampilkan pada Gambar 9.

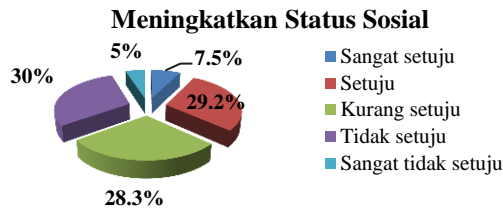
### Pekerjaan Ayah



**Gambar 10.** Diagram variabel Pekerjaan Ayah

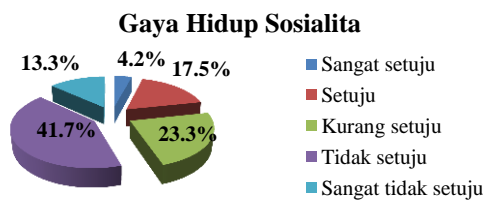
Pada Gambar 10 menunjukkan bahwa Pekerjaan Ayah responden penelitian yang paling banyak adalah sebagai PNS yaitu sebanyak 31 orang atau

25.8% sedangkan pekerjaan Ayah dengan kategori lain seperti pegawai swasta sebanyak 11 orang atau 9.2%, wiraswasta sebanyak 19 orang atau 15.8%, petani 27 orang atau 22.5% dan yang pekerjaan ayahnya di luar kategori yang ada sebanyak 32 orang atau 26.7%.



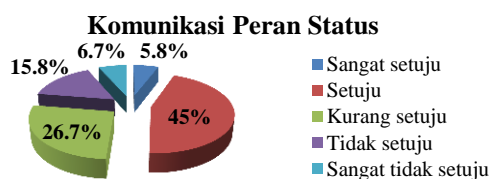
**Gambar 11.** Diagram Distribusi variabel Tambahan Meningkatkan Status Sosial

Pada Gambar 11 menunjukkan jawaban responden untuk variabel pertanyaan mengenai apakah membeli pakaian dapat meningkatkan status sosial responden penelitian, dimana yang paling dominan menjawab dengan jawaban tidak setuju yaitu sebanyak 36 orang atau 30%.



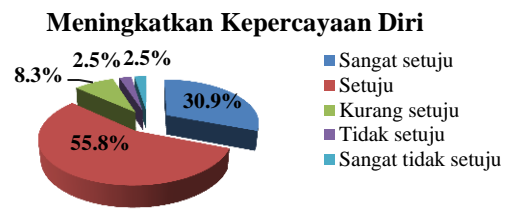
**Gambar 12.** Diagram Distribusi variabel Tambahan Gaya Hidup Sosialita

Pada Gambar 12 menunjukkan jawaban responden untuk variabel pertanyaan mengenai apakah membeli pakaian untuk mengikuti gaya hidup sosialita tertentu, dimana yang paling dominan menjawab dengan jawaban tidak setuju yaitu sebanyak 50 orang atau 41.7% sedangkan ntuk penyebaran jawaban lain dapat dilihat pada diagram pada Gambar 12.



**Gambar 13.** Diagram Distribusi variabel Tambahan Komunikasi Peran Status

Pada Gambar 13 menunjukkan jawaban responden untuk variabel pertanyaan mengenai apakah membeli pakaian dapat meningkatkan Komunikasi Peran Status responden penelitian, dimana yang paling dominan menjawab dengan jawaban tidak setuju yaitu sebanyak 54 orang atau 45% sedangkan untuk penyebaran jawaban lain dapat dilihat pada diagram pada Gambar 13.



**Gambar 14.** Diagram Distribusi variabel Tambahan Meningkatkan Kepercayaan Diri

Pada Gambar 14 menunjukkan frekuensi dan Persentase jawaban responden untuk variabel tambahan. Pertanyaan mengenai apakah anda setuju bahwa membeli pakaian dapat meningkatkan Kepercayaan Diri? Jawaban responden didominasi dengan pernyataan setuju sebanyak 67 orang atau 55.8%.

**Tabel 3.** Rekapitulasi data variabel terikat Y

Variabel Y	Pengeluaran konsumsi pakaian
Minimum	0.0
Maksimum	2000
Rata-rata	107.7
Varian	100.0

Output pada Tabel 3 mendeskripsikan nilai minimum, median, rata-rata dan nilai maksimum pada variabel terikat. Misalkan pada variabel terikatnya, menunjukkan nilai minimumnya adalah nol dimana nilai tersebut adalah sensor variabel Y, kemudian nilai maksimum untuk variabel terikat adalah 2000 yang menunjukkan bahwa pengeluaran biaya konsumsi responden paling besar adalah Rp.2.000.000.

#### D. Pengujian Model Regresi Tobit Standar

##### 1. Estimasi Parameter

Hasil dari listing untuk model regresi tobit berupa observasi variabel yang meliputi total observasi, sensor kiri, sensor kanan dan tidak tersensor yang diringkas pada tabel berikut:

**Tabel 4.** Observasi variabel sensor

Sensor Kiri	40 observasi
Sensor Kanan	0 observasi
Tidak tersensor	80 observasi
Total Observasi	120 observasi

Sensor Kiri yang dimaksud dalam Tabel 4 adalah data pada kuesioner hasil penelitian terkonsentrasi atau terkumpul pada titik nol. Dari total 120 observasi terdapat 40 responden yang mengisi pengeluaran konsumsi sebesar nol rupiah dan 80 responden lainnya mengisi dengan nilai yang bervariasi yaitu lebih dari nol rupiah dan kurang dari sama dengan 2 juta rupiah, sehingga tergolong kedalam observasi yang tidak tersensor.

**Tabel 5.** Estimasi dugaan model awal

Koefisien	Estimasi	Standar Error
Intercept (1)	-13.052	258.871
Jenis Kelamin ( $X_1$ )	-115.189	86.639
Asal Daerah ( $X_2$ )	-295.575	106.019
Uang Sumber lain selain dari tua ( $X_3$ )	0.0388	0.058
Penghasilan tua ( $X_4$ )	2.703	38.828
Tempat Tinggal ( $X_5$ )	15.345	66.903
Jumlah Saudara ( $X_6$ )	11.117	22.134
Jumlah Uang Bulanan ( $X_7$ )	0.605	0.182
Pekerjaan Ayah ( $X_8$ )	14.986	26.830
Status Sosial ( $X_9$ )	-7.699	49.576
Gaya hidup Sosialita ( $X_{10}$ )	47.549	43.968
Komunikasi peran status ( $X_{11}$ )	63.375	44.049
meningkatkan kepercayaan diri ( $X_{12}$ )	-123.909	58.309

Berdasarkan output *coefficients* terdapat variabel *intercept* yang berarti akan terdapat satu model dugaan awal dengan masing-masing nilai dugaan, standar error dan nilai z hitung, demikian pula dengan 12 variabel bebasnya. Berdasarkan model matematisnya pada Persamaan (3), didapatkan model tobit standarnya adalah sebagai berikut:

$$\hat{y}^* = -13.05 - 115.19X_1 - 295.56X_2 + 0.04X_3 + 2.70X_4 + 15.34X_5 + 11.12X_6 + 0.61X_7 + 14.98X_8 - 7.70X_9 + 47.55X_{10} + 63.37X_{11} - 123.91X_{12}$$

Dimana

$$\hat{y} = \begin{cases} \hat{y}^* & \text{jika } \hat{y}^* > 0 \\ 0 & \text{jika } \hat{y}^* \leq 0 \end{cases}$$

Coeffisien:	Estimasi	Standar Error	t value	Pr(>t)
(Intercept)	-14.373	251.831	-0.057	0.954
X1	-111.178	86.654	-1.283	0.199
X2	-289.919	101.371	-2.86	0.004
X3	0.036	0.057	0.62	0.536
X4	2.757	37.923	0.073	0.942
X5	16.194	65.341	0.248	0.804
X6	11.680	21.313	0.548	0.584
X7	0.592	0.180	3.288	0.001
X8	14.262	26.127	0.546	0.585
X9	-6.948	48.162	-0.144	0.885
X10	45.918	42.657	1.076	0.282
X11	62.407	43.864	1.423	0.155
X12	-121.083	57.264	-2.114	0.034
logSigma	5.897	0.083	70.636	< 2e-16

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Newton-Raphson maximisation, 10 iterations  
Return code 1: gradient close to zero  
Log-likelihood: -614.8203 on 14 Df

**Gambar 15.** Output *log Sigma* uji serentak

Berdasarkan Gambar 15, pada kotak yang berwarna merah menunjukkan nilai *logSigma* yang mana *logSigma* tersebut digunakan sebagai uji serentak pada model regresi tobit. Dari pengujian tersebut diperoleh nilai Probabilitas untuk *logSigma* adalah 0.000 (2e-16), sedangkan nilai  $\alpha = 0.05$  yang berarti nilai  $Prob < \alpha$  atau  $0.000 > 0.05$  dengan kata lain hipotesis nol ditolak yang artinya setidaknya ada satu variabel yang berpengaruh atau signifikan di dalam model.

## 2. Uji Individu menggunakan Uji Wald

Berdasarkan Tabel 6, nilai *p-value* yang kurang dari  $\alpha = 0.05$  dapat disimpulkan sebagai variabel

yang signifikan. Dari output tersebut tampak ada beberapa variabel yang nilai *p-value*nya kurang dari  $\alpha = 0.05$  yaitu Asal Daerah ( $X_2$ ), jumlah uang bulanan ( $X_7$ ) dan alasan seseorang membeli pakaian untuk meningkatkan kepercayaan diri ( $X_{12}$ ) yaitu masing-masing 0.00554, 0.000989 dan 0.032468.

Oleh karena itu dapat disimpulkan bahwa pada pengujian secara individu hanya diperoleh tiga variabel yang secara signifikan mempengaruhi variabel terikat *Y* pada tingkat kepercayaan 95%.

**Tabel 6.** Uji Individu variabel *X* dengan variabel *Y* dengan keputusan

Variabel	Estimasi	Wald	P-value	Keputusan
Jenis Kelamin ( $X_1$ )	-115.189	-1.322	0.188	H <sub>0</sub> diterima
Asal Daerah ( $X_2$ )	-295.575	-2.801	0.006	H <sub>0</sub> ditolak
Uang dari Sumber Lain ( $X_3$ )	0.039	0.685	0.494	H <sub>0</sub> diterima
Penghasilan Tua ( $X_4$ )	2.703	0.070	0.944	H <sub>0</sub> diterima
Tempat Tinggal ( $X_5$ )	15.345	0.228	0.820	H <sub>0</sub> diterima
Jumlah Saudara ( $X_6$ )	11.117	0.503	0.615	H <sub>0</sub> diterima
Jumlah Uang Bulanan ( $X_7$ )	0.605	3.337	0.001	H <sub>0</sub> ditolak
Pekerjaan Ayah ( $X_8$ )	14.986	0.554	0.580	H <sub>0</sub> diterima
Meningkatkan Status Sosial ( $X_9$ )	-7.699	-0.155	0.877	H <sub>0</sub> diterima
Gaya Hidup Sosialita ( $X_{10}$ )	47.549	1.081	0.281	H <sub>0</sub> diterima
Komunikasi peran status ( $X_{11}$ )	63.375	1.428	0.155	H <sub>0</sub> diterima
Meningkatkan kepercayaan diri ( $X_{12}$ )	-123.909	-2.152	0.032	H <sub>0</sub> ditolak

## 3. Model Terbaik Regresi Tobit Standar

Dilihat berdasarkan hasil yang terdapat dalam Tabel 21, yang menunjukkan signifikansi variabel-variabel bebas dapat dilihat berdasarkan nilai Wald yang dibandingkan dengan mutlak *z* – Tabel atau dapat dilihat dari nilai *p-value* yang dibandingkan dengan nilai  $\alpha = 0.05$ . Berdasarkan hasil tersebut dapat disimpulkan bahwa model terbaik regresi tobit standar adalah sebagai berikut:

$$y = \beta_0 + \beta_1 X_2 + \beta_2 X_7 + \beta_3 X_{12}$$

$$y = -13.0521 - 295.575X_2 + 0.605X_7 - 123.909X_{12}$$

## 4. Interpretasi Parameter Model

Dari model yang diperoleh dapat diartikan bahwa pengeluaran konsumsi pakaian dikalangan Mahasiswa Matematika Angkatan 2010-2013 FMIPA Unib dipengaruhi oleh asal daerah responden. Berdasarkan angka yang diperoleh menunjukkan bahwa variabel asal daerah bernilai  $-295.56$ , artinya ada kemungkinan responden yang berasal dari dalam Provinsi Bengkulu lebih menunjukkan pengaruh terhadap biaya pengeluaran konsumsi pakaian dikalangan mahasiswa Matematika Angkatan 2010-2013 FMIPA Unib dibandingkan dengan yang berasal dari luar Provinsi Bengkulu.

Selanjutnya untuk variabel kedua yang dinyatakan berpengaruh terhadap biaya pengeluaran konsumsi pakaian dikalangan mahasiswa adalah variabel jumlah uang bulanan ( $X_7$ ). Nilai variabel jumlah uang bulanan yang bernilai 0.61 dapat berarti bahwa semakin besar jumlah uang bulanan maka akan cukup berpengaruh terhadap biaya pengeluaran konsumsi pakaian dikalangan mahasiswa.

Pada variabel yang dinyatakan mempunyai pengaruh terhadap biaya pengeluaran konsumsi selain dua variabel di atas adalah variabel pertanyaan mengenai apakah responden setuju dengan pernyataan bahwa membeli pakaian dapat meningkatkan kepercayaan diri? Namun pada variabel penelitian ini, seharusnya antara variabel ordinal, nominal dan rasio tidak dapat dilakukan dalam satu pengujian yang sama, sehingga mengakibatkan terjadi *abuse* atau penyalagunaan yang tidak disarankan dalam suatu penelitian.

Dari hasil penelitian ini dapat dilihat nilai dari variabel ( $X_{12}$ ) yaitu  $-123.91$  yang dapat diartikan bahwa jawaban responden dominan menjawab pada kategori setuju, sehingga dari model dapat ditarik indikasi bahwa alasan responden mengeluarkan biaya konsumsi pakaian agar dapat meningkatkan kepercayaan diri.

### 5. Uji Kelayakan Model

Hasil pengujian kelayakan model regresi tobit standar untuk kasus pengeluaran konsumsi pakaian dikalangan mahasiswa terdapat tampak bahwa nilai  $R^2$  yang diperoleh tergolong sangat kecil yaitu  $R^2 = 0.2194$ . Nilai kebaikan model ini memiliki arti bahwa dari dua belas variabel bebas yang digunakan dalam penelitian ini, hanya dapat menjelaskan variabel pengeluaran konsumsi pakaian dikalangan mahasiswa Matematika angkatan 2010-2013 FMIPA Unib sebesar 22% sedangkan 78% lainnya dipengaruhi oleh faktor-faktor lain yang tidak teridentifikasi pada penelitian ini. Hal tersebut terjadi dimungkinkan karena sampel yang digunakan terlalu kecil dan variabel yang digunakan terlalu banyak.

### E. Uji Sifat-sifat Dasar Model Tobit Standar

#### 1. Uji Kenormalan Tersensor Variabel Terikat

Akan diuji terlebih dahulu apakah  $Y$  berdistribusi normal atau tidak.

**Tabel 7.** Uji Kenormalan dengan *Kolmogorov-Smirnov Test*

		Konsumsi
Jumlah variabel		120
Parameter Normal	Rataan	187.740
	Standar Deviasi	311.404
Kolmogorov-Smirnov Z		2.994
Asymp. Sig. (2-tailed)		0.000

Berdasarkan uji Kolmogorv-Smirnov di atas, dengan Hipotesis:

$H_0$  : Data menyebar normal

$H_1$  : Data tidak menyebar normal

Berdasarkan statistik uji yang digunakan untuk mengetahui kenormalan data maka pada output dapat dilihat bahwa nilai Asymp. Sig (2-Tailed) yaitu 0.000 yang lebih kecil daripada nilai  $\alpha = 0.05$

sehingga dapat disimpulkan bahwa  $H_0$  ditolak yang berarti data tidak menyebar secara normal.

### 2. Uji Multikolinieritas

Pada pengujian asumsi multikolinieritas digunakan program SPSS 16. Suatu model regresi dikatakan terjadi multikolinieritas jika mempunyai nilai *Variance Inflation factor* (VIF) lebih dari 10 atau nilai *Tolerance* tidak sama dengan 1.

**Tabel 8.** Diagnosa Kolinieritas

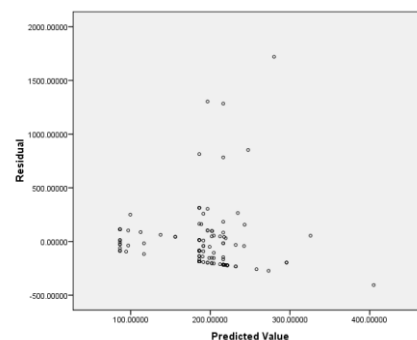
Model	Statistik Kolinieritas	
	Tolerance	VIF
JK	0.891	1.122
Asal Daerah	0.782	1.279
Uang Sumber Lain	0.939	1.065
Penghasilan Orang tua	0.584	1.712
Tempat Tinggal	0.739	1.354
Jumlah Saudara	0.888	1.127
Uang Bulanan	0.497	2.010
Pekerjaan Ayah	0.765	1.307
Status Sosial	0.510	1.960
Gaya Hidup Sosialita	0.639	1.565
Komunikasi Peran Status	0.649	1.541
Meningkatkan Kepercayaan Diri	0.618	1.618

a. Variabel Terikat: Konsumsi

Dari Tabel 8 diketahui nilai pada variabel mempunyai angka VIF kurang dari 10 dan nilai Tolerance mendekati 1 yaitu untuk masing-masing variabel bebas sehingga dapat disimpulkan model regresi tidak memenuhi asumsi multikolinieritas. Dengan demikian tidak terjadi hubungan linier yang sempurna antar variabel bebas.

### 3. Uji Heteroskedastisitas

Pada pengujian asumsi heteroskedastisitas digunakan program SPSS 16. Pada Gambar 15 tampak sisaan menyebar tidak merata di atas dan di bawah sumbu  $Y$  dengan kata lain penyebaran data antara nilai *predicted value* dan residual mengikuti suatu pola tertentu yang berarti model regresi memenuhi asumsi heteroskedastisitas.



**Gambar 15.** Grafik *Scatterplot*

Untuk lebih memastikan dapat menggunakan uji F. Berdasarkan Tabel 9 dapat dilihat bahwa nilai  $F_{Hitung} = 2.506$  sedangkan nilai  $F_{Tabel} = 2.341$  yang berarti nilai  $F_{Hitung} > F_{Tabel}$  sehingga dapat ditarik kesimpulan bahwa  $H_0$  ditolak yang berarti ragam error bersifat heteroskedastisitas.

**Tabel 9.** Uji F

Model	Jumlah Kuadrat	db	Rataan Kuadrat	F	Sig.
1 Regresi	2531923.086	12	210993.591	2.506	0.006
Error	9007775.905	107	84184.822		
Total	1.154E7	119			

**4. Uji Autokorelasi**

Asumsi ini dapat dilihat dengan uji Durbin-Watson yang akan dibandingkan dengan batas atas atau batas bawah nilai tabel Durbin-Watson.

**Tabel 10.** Uji Durbin-Watson

Model	R	R Kuadrat	R Kuadrat Biasa	Standar Error estimasi	Durbin-Watson
1	0.468 <sup>a</sup>	0.219	0.132	290.195	0.563

a. Predictors: (Constant), Meningkatkan Kepercayaan Diri, Jenis Kelamin, Asal Daerah, Jumlah Saudara, Penghasilan Orang tua, Uang Sumber Lain, Gaya Hidup Sosialita, Tempat Tinggal, Pekerjaan Ayah, Komunikasi Peran Status, Status Sosial, Uang Bulanan

Hipotesis nol ditolak karena nilai  $0 < d < d_L$  yang dapat dilihat pada Tabel 1. Dari pengujian residual pada 12 variabel dan  $n = 120$  maka diperoleh nilai Durbin-Watson untuk hasil pengujiannya adalah  $d = 0.563$  dan nilai  $d_L$  dan  $d_u$  pada tabel Durbin-Watson adalah 1.4889 dan 1.9256, maka dari sana dapat disimpulkan bahwa Hipotesis nol ditolak artinya adanya autokorelasi positif pada data tersebut.

**V. PENUTUP**

**A. Kesimpulan**

Berdasarkan hasil penelitian dan pengujian dari bab sebelumnya, diperoleh kesimpulan sebagai berikut:

1. Dari 12 variabel yang diuji, variabel yang dinyatakan signifikan atau mempengaruhi adalah variabel Asal Daerah ( $X_2$ ), Jumlah Uang Bulanan ( $X_7$ ) dan variabel pertanyaan mengenai apakah membeli pakaian dapat meningkatkan kepercayaan diri ( $X_{12}$ ). Sehingga model yang terbentuk adalah:

$$y = -13.0521 - 295.575X_2 + 0.605X_7 - 123.909X_{12}$$

Sedangkan untuk variabel lain yang mempengaruhi pengeluaran biaya konsumsi mahasiswa Matematika FMIPA Unib disebabkan oleh faktor lain diluar variabel-variabel yang diuji.

2. Berdasarkan hasil pengujian kebaikan kesesuaian model dengan determinasi  $R^2$  diperoleh hasil yang tidak terlalu baik yaitu  $R^2 = 0.2194$  atau kebaikan model sebesar 22%.

**B. Saran**

Bila menganalisis suatu masalah yang mengandung variabel dependen yang tersensor atau terkelompok pada nilai batas nol (0) dapat digunakan Model Tobit Standar maupun variasi dari model tobit

lainnya. Sebagai referensi pengembangan penelitian selanjutnya pembaca dapat menganalisis regresi tobit untuk beberapa model tobit diperumum lainnya (*Generalized Tobit Model*).

**DAFTAR PUSTAKA**

Amemiya, T. 1985. *Advanced Econometrics*. Harvard University Press. Cambridge, Massachusetts. USA.

Anwar, K. 2011. *Analisis Pola Konsumsi Masyarakat Pedesaan di Kabupaten Bireuen-Aceh*. Fakultas Ekonomi. Universitas Malikussaleh. Aceh.

Arikunto, S. 2006. *Prosedur penelitian: suatu pendekatan praktek*. Rineka Cipta. Jakarta.

Bierens, H. J. 2004. The Tobit Model. <URL: <http://econ.la.psu.edu/~Tobit.PDF/html>> Diakses tanggal 17 April 2014

Frone, R. M. 1997, *Regression Models For Discrete and Limited Dependent Variabels*, <URL:[http://division.aomonline.org/rm/1997\\_forum\\_regression\\_models.htmlfrone](http://division.aomonline.org/rm/1997_forum_regression_models.htmlfrone)>.Diakses tanggal 24 April 2014.

Greene, W. H. 2008. *Econometrics Analysis, 6<sup>th</sup> Edition*. New Jersey: Prentice Hall.

Gujarati, D. 2004. *Basic Econometrics. Fourth Edition*. International Edition. Singapore. McGraw-Hill.

Hosmer, D. W., and S. Lemeshow. 2000. *Applied Logistic Regression, Second Edition*. John Willey. New York.

Joreskog, K. G. 2002. *Censored Variabels and Censored Regression*, <http://aac.asm.org/cgi/reprint/50/1/62> . Diakses tanggal 24 April 2014.

Kutner, M. H., C. J. Nachtsheim J. Neter and W. Li. 2004. *Applied Linear Regression Models*. New York: McGraw-Hill/Irwin.

Larissa, I. D. dan D. Ispriyanti. 2008. Penentuan Estimasi Parameter Regresi dengan Variabel Dependen Tersensor. *Jurnal Matematika* 11(3): 135-140, ISSN: 1410

Long, J. S. 2001. *Regression Models for Categorical Dependent Variables Using Stata*. College Station. Texas.

Olsen, R. 1978. "Note on the Uniqueness of the Maximum Likelihood Estimator for the Tobit Model". *Econometrica*. Vol. 46. No. 5

Permana, G. 2011. *Analisis Regresi Tobit Pada Permasalahan Pengeluaran Konsumsi Rokok Kota Kediri Tahun 2011*. Jurusan Matematika. FMIPA. Universitas Brawijaya. Malang. <URL: <http://statistik.studentjournal.ub.ac.id/index.php/statistik/article/download/72/74> >. Diakses pada: tanggal 06 Mei 2014.

Salim, T. 2007. Model Tobit Standar dan Model Heckit. [SKRIPSI]. Fakultas MIPA. UI: Depok. <URL: <http://lontar.ui.ac.id/file?file=digital/20180918-011-07-Model%20Tobit.pdf> >. Diakses pada: tanggal 16 April 2014.

Widarjono, A. 2007. *Ekonometrika: Teori dan Aplikasi untuk Ekonomi dan Bisnis*. Edisi Kedua. Yogyakarta: Ekonisia Fakultas Ekonomi Universitas Islam Indonesia.

Wulandari, S.P., M. Salamah dan D. Susilaningrum. 2009. *Diktat Pengajaran Analisis Data Kualitatif*. Surabaya.

# ANALISIS TINGKAT POTENSI KEBANGKRUTAN PADA PERUSAHAAN SEKTOR PROPERTI DAN *REAL ESTATE* DENGAN MENGGUNAKAN REGRESI LOGISTIK BINER DAN ANALISIS DISKRIMINAN

<sup>1</sup>Dori Hertati Hutapea, <sup>2</sup>Sigit Nugroho, <sup>3</sup>Etis Sunandi

<sup>1</sup>Mahasiswa Jurusan Matematika FMIPA Universitas Bengkulu

Email : [dorihertatihutapea10@gmail.com](mailto:dorihertatihutapea10@gmail.com)

<sup>2,3</sup>Staff Pengajar Jurusan Matematika FMIPA Universitas Bengkulu

## ABSTRAK

Regresi logistik biner merupakan persamaan yang menghasilkan nilai peluang yang dipakai sebagai dasar untuk klasifikasi. Analisis regresi logistik biner tidak memerlukan uji asumsi seperti halnya analisis diskriminan, sehingga analisis regresi logistik biner dapat diterapkan dalam berbagai skala data. Pada analisis diskriminan diperlukan asumsi multivariat normal, terdapat perbedaan rata-rata antara dua kelompok data, tidak ada multikolinieritas dan matriks korelasi kedua kelompok data sama atau homogen. Hasil penelitian ini menunjukkan bahwa variabel yang signifikan berpengaruh terhadap tingkat potensi kebangkrutan perusahaan properti dan *real estate* tahun 2012-2016 dengan taraf pengujian 5% pada analisis diskriminan yaitu variabel  $\ln DER$  dan  $\sqrt{GPM}$  sedangkan pada regresi logistik biner yaitu variabel NPM. Oleh karena variabel signifikan yang dihasilkan dari kedua analisis tersebut berbeda maka hasil uji kedua analisis tidak dapat dibandingkan. Namun, jika hasil perbandingan kedua analisis tersebut dengan menggunakan variabel yang signifikan dan variabel yang tidak signifikan dimasukkan kedalam model, dapat disimpulkan bahwa hasil analisis regresi logistik biner memiliki rata-rata ketepatan prediksi klasifikasi lebih besar dibandingkan dengan analisis diskriminan.

**Kata kunci:** Kebangkrutan, Regresi, Logistik Biner, Diskriminan, Klasifikasi.

## 1. PENDAHULUAN

Pertumbuhan penjualan sektor properti Indonesia bertumbuh cepat pada tahun 2012 hingga 2013, sejak bulan Juli 2012 hingga April 2013 penjualan properti untuk daerah Depok setiap bulan meningkat, diakhir bulan April 2013 total properti yang dipasarkan mencapai 1602 properti yang meliputi rumah sebanyak 1157 buah, tanah sebanyak 361 unit, ruko dan ruangan usaha atau property komersial sebanyak 84 buah (Claradiesty, 2016). Pada tahun 2012 dan pertengahan pertama tahun 2013 pertumbuhan keuntungan properti Indonesia naik sebesar 26 perusahaan mencatat pertumbuhan laba bersihnya yang lebih dari 50% mengakibatkan harga properti Indonesia meningkat. Kondisi tersebut menjadikan sektor properti kembali meningkat (Anonim, 2015). Pada tahun 2015 kuartal I penjualan properti mengalami penurunan yang sangat signifikan. Penurunan

terbesar untuk pertumbuhan properti dirasakan di pasar wilayah Surabaya, Jakarta Bogor Depok Tangerang & Bekasi (Jabodetabek), serta Manado. Secara tahunan, harga properti residensial juga mengalami kenaikan yang melambat. Penjualan properti kuartal II tahun 2016 mulai meningkat kembali menjadi 4,02% (Azkia, 2016).

Penurunan pertumbuhan penjualan akan menyebabkan penurunan pendapatan perusahaan-perusahaan yang bergerak disektor properti dan *real estate* yang secara langsung akan berdampak kepada penurunan laba atau profitabilitas perusahaan (Kusumajaya, 2011). Selain berdampak kepada laba, penurunan penjualan akan berdampak terhadap struktur hutang dari perusahaan tersebut. Sebuah perusahaan apabila tingkat penjualan atau pertumbuhan penjualan dari tahun ke tahun memiliki pertumbuhan yang stabil dan dapat menghasilkan profitabilitas yang tinggi akan membuat perusahaan lebih mengutamakan pendanaan secara internal (laba) daripada eksternal (hutang) untuk kegiatan operasional maupun untuk membayar kewajibannya (Amalia, 2011). Hal itu dilakukan guna menghindari akan adanya resiko gagal bayar hingga resiko kebangkrutan. Resiko kebangkrutan perusahaan terlihat dari kondisi keuangan perusahaan. Kondisi keuangan perusahaan biasanya tersaji dan diinformasikan melalui laporan keuangan yang berupa neraca dan laporan rugi laba. Laporan keuangan perusahaan terdapat rasio-rasio keuangan, diantaranya rasio likuiditas, rasio *leverage*, rasio aktivitas, rasio probabilitas dan rasio nilai pasar.

Ada banyak metode yang digunakan dalam analisis tingkat potensi kebangkrutan. Analisis tersebut bertujuan untuk mengklasifikasi variabel respon ke dalam dua grup atau dua kelompok. Beberapa alat analisis penelitian statistika yang dapat digunakan dalam analisis pengklasifikasi tingkat potensi perusahaan yaitu, analisis Regresi Logistik Biner dan Analisis Diskriminan. Metode Regresi Logistik Biner dan Analisis Diskriminan memiliki kegunaan yang sama. Kegunaan kedua metode tersebut adalah pengklasifikasian, sehingga dapat melihat hasil perbandingan klasifikasi antara regresi logistik biner dan analisis diskriminan. Salah satu penerapannya, peneliti melakukan penelitian dengan mengambil judul “**Analisis Tingkat Potensi Kebangkrutan pada Perusahaan Sektor Properti dan *Real Estate* dengan menggunakan Regresi Logistik Biner dan Analisis Diskriminan**”. Tujuan dari penelitian adalah dapat menerapkan prosedur Analisis Regresi Logistik dan Analisis Diskriminan Biner dalam menganalisis tingkat potensi kebangkrutan Perusahaan Sektor Properti dan *Real Estate* pada BEI serta membandingkan hasil ketepatan klasifikasi tingkat potensi kebangkrutan pada Perusahaan Sektor Properti dan *Real Estate*.

## 2. METODE PENELITIAN

### 2.1 Jenis Data

Jenis data dari penelitian ini adalah data sekunder, yaitu data tingkat potensi kebangkrutan perusahaan sektor properti dan *real estate* yang terdaftar pada Bursa Efek Indonesia (BEI) Tahun 2012-2016

### 2.2 Variabel Penelitian

Variabel yang digunakan pada penelitian ini laporan tahunan yang dipublikasikan pada Bursa Efek Indonesia dari tahun 2012 s/d 2016 yang meliputi *Current Ratio* (CR), *Quick Ratio* (QR), *Debt to Asset Ratio* (DAR), *Debt to Equity Ratio* (DER), *Return On Asset* (ROA), *Return On Equity* (ROE), *Gross Profit Margin* (GPM), *Operating Profit Margin* (OPM), *Net Profit Margin* (NPM). Data yang digunakan merupakan data tahunan dengan periode 2012 s/d 2016 yang berjumlah 170 observasi.

### 2.3. Teknik Analisis Data

Langkah-langkah analisis yang digunakan pada penelitian ini adalah sebagai berikut:



Gambar 1. Flowchart Analisis Regresi Logistik Biner dan Analisis Diskriminan.

## 3. HASIL DAN PEMBAHASAN

### 3.1 Analisis Data Deskriptif

Sampel pada penelitian ini berjumlah 34 perusahaan sektor Properti dan *Real Estate* yang terdaftar di Bursa Efek Indonesia dengan rentang waktu pengambilan data sampel dari tahun 2012-2016. Data yang digunakan untuk analisis ini, dapat dilihat pada Tabel 1.

**Tabel 1. Jumlah Perusahaan Pada Masing-Masing Grup**

Kelompok	Jumlah Perusahaan
Financial Distress atau Bangkrut	5
Non Financial Distress atau Tidak Bangkrut	29
<b>Total</b>	<b>34</b>

### 3.2 Uji Asumsi Diskriminan

#### 3.2.1 Uji *Multivariate Normality*

Hipotesis yang berlaku adalah  $H_0$ : Variabel prediktor berdistribusi Normal vs  $H_1$ : Variabel prediktor tidak berdistribusi Normal. Taraf nyata pengujian  $\alpha = 5\%$ . Hasil Uji *Multivariate Normal* sebagai berikut:

**Tabel 2 Hasil *One-Sample Kolmogorov-Smirnov Test***

Data Asli			Data Transformasi		
Variabel Prediktor		Signifikan	Variabel Prediktor		Signifikan
$X_1$	CR	0.000	$X_1^*$	$\frac{1}{CR}$	0.171
$X_2$	QR	0.000	$X_2^*$	$\ln QR$	0.105
$X_3$	DAR	0.188	$X_3$	DAR	0.188
$X_4$	DER	0.018	$X_4^*$	$\ln DER$	0.249
$X_5$	ROA	0.047	$X_5^*$	$\ln ROA$	0.674
$X_6$	ROE	0.250	$X_6$	ROE	0.250
$X_7$	GPM	0.0026	$X_7^*$	$\sqrt{GPM}$	0.065
$X_8$	OPM	0.000	$X_8^*$	$\ln OPM$	0.066
$X_9$	NPM	0.000	$X_9^*$	$\ln NPM$	0.149

Kriteria pengambilan keputusan uji *Multivariate Normality* adalah jika nilai signifikansi suatu variabel  $> 0.05$  maka data tersebut berdistribusi normal. Berdasarkan Tabel 2, menunjukkan bahwa hanya variabel  $X_3$  dan  $X_6$  terdistribusi dengan normal. Oleh karena itu, variabel  $X_1, X_2, X_4, X_5, X_7, X_8$  dan  $X_9$  harus dilakukan transformasi data sehingga variabel tersebut terdistribusi dengan normal.

#### 3.2.2 Uji Vektor Rata-Rata Kedua Kelompok (*Independent Sample T-Test*)

Uji *Independent Sample T-Test* dilakukan pada variabel-variabel yang memiliki data terdistribusi dengan normal. Hipotesis yang digunakan, yaitu  $H_0$  : tidak terdapat perbedaan rata-rata antara dua kategori vs  $H_1$  : terdapat perbedaan rata-rata antara dua kategori

**Tabel 3. Hasil Uji *Independent Sample Test***

Rasio	Signifikasi	
$X_1^*$	$\frac{1}{CR}$	0.000
$X_2^*$	$\ln QR$	0.000
$X_3$	DAR	0.000
$X_4^*$	$\ln DER$	0.000
$X_5^*$	$\ln ROA$	0.000
$X_6$	$\ln ROE$	0.000

**Lanjutan Tabel 3...**

$X_7^*$	$\sqrt{GPM}$	0.000
$X_8^*$	$\ln OPM$	0.009
$X_9^*$	$\ln NPM$	0.007

Berdasarkan Tabel 3, dapat dilihat bahwa masing-masing variabel prediktor / independen yang digunakan mempunyai nilai signifikan  $< 0.05$  yang berarti kesembilan variabel tersebut terdapat perbedaan rata-rata antara dua kategori

**3.2.3 Uji Multikolinearitas**

Hipotesis yang berlaku adalah  $H_0$ : jika nilai VIF  $< 10$ , maka tidak terjadi *multikolinearitas* diantara variabel predictor / independen vs  $H_1$ : jika nilai VIF  $> 10$ , maka terjadi *multikolinearitas* diantara variabel prediktor / independen

**Tabel 4 Hasil Uji Multikolinearitas**

Data 1			Data 2		
Variabel Prediktor		VIF	Variabel Prediktor		VIF
$X_1^*$	$\frac{1}{CR}$	112.727	$X_4^*$	$\ln DER$	1.026
$X_2^*$	$\ln QR$	5.235	$X_5^*$	$\ln ROA$	2.559
$X_3$	$DAR$	168.228	$X_7^*$	$\sqrt{GPM}$	1.312
$X_4^*$	$\ln DER$	64.719	$X_8^*$	$\ln OPM$	2.383
$X_5^*$	$\ln ROA$	8.743	$X_9^*$	$\ln NPM$	3.853
$X_6$	$ROE$	5.740			
$X_7^*$	$\sqrt{GPM}$	1.596			
$X_8^*$	$\ln OPM$	2.485			
$X_9^*$	$\ln NPM$	5.049			

Berdasarkan Tabel 4, menunjukkan bahwa hasil perhitungan nilai VIF pada data 1 menunjukkan ada variabel yang mempunyai nilai VIF  $> 10$ , sedangkan pada data 2 menunjukkan bahwa tidak ada variabel prediktor dan independen yang mempunyai nilai VIF  $> 10$ .

**3.2.4 Uji Kesamaan Matrik Korelasi**

Hipotesisi yang berlaku adalah  $H_0$ : Matriks korelasi kedua kelompok adalah sama atau homogen vs  $H_1$ : Matriks korelasi kedua kelompok adalah tidak sama atau heterogen

**Tabel 5. Uji Kesamaan Matriks Korelasi**

<b>Box's M</b>		7.042
<b>F</b>	<b>Approx.</b>	2.275
	<b>df1</b>	3
	<b>df2</b>	23069.385
	<b>Sig.</b>	0.078

Berdasarkan Tabel 5, menunjukkan bahwa nilai *Sig. Test Results* statistik uji *Box' M* sebesar 0.078 yang berarti  $> 0.05$  maka  $H_0$  diterima. Dengan demikian matriks korelasi kedua kelompok data sama atau homogen.

### 3.3 Analisis Diskriminan

Berdasarkan pemilihan variabel diskriminan dengan menggunakan metode *stepwise* maka diperoleh dua variabel yaitu variabel rasio keuangan  $\ln DER (X_4^*)$  dan  $\sqrt{GPM}(X_7^*)$  yang mempunyai kekuatan terbesar dalam diskriminan. Model diskriminan yang terbentuk dengan pemilihan variabel diskriminan, yaitu:

$$Z \text{ score} = -1.203 + 1.190X_4^* + 0.001 X_7^*$$

Dimana :

$Z \text{ score}$  = *Score Diskriminant*

$X_4^*$  = *ln DER (ln Debt to Equity Ratio)*

$X_7^*$  =  $\sqrt{GPM}$  (*√Gross Profit Margin*)

Fungsi ini berguna untuk menganalisis observasi yang diteliti akan termasuk ke dalam kelompok pertama (keputusan 0 yaitu potensi perusahaan bangkrut) atau kedua (keputusan 1 yaitu potensi perusahaan tidak bangkrut).

Sedangkan model diskriminan yang terbentuk dengan menggunakan kesembilan variabel dimasukkan kedalam formula persamaan fungsi diskriminan, adalah sebagai berikut:

$$Z \text{ score} = 2.979 + 9.849X_1^* - 0.241X_2^* - 17.763X_3^* + 2.263X_4^* - 0.955X_5^* \\ + 0.101X_6^* + 0.000X_7^* - 0.120X_8^* + 0.515X_9^*$$

Dimana :

$Z \text{ score}$  = *Score Diskriminant*

$X_1^*$  =  $\frac{1}{\text{Current Ratio}}$

$X_2^*$  = *ln Quick Ratio*

$X_3^*$  = *Debt to Asset Ratio*

$X_4^*$  = *ln Debt to Equity Ratio*

$X_5^*$  = *ln Return On Asset*

$X_6^*$  = *ln Return On Equity*

$X_7^*$  =  $\sqrt{\text{Gross Profit Margin}}$

$X_8^*$  = *ln Operating Profit Margin*

$X_9^*$  = *ln Net Profit Margin*

Model analisis diskriminan menunjukkan bahwa rata-rata ketepatan prediksi AD model 1 sebesar 82.75%, peluang kesalahan klasifikasi sebesar 0.1529, presentase ketepatan pengklasifikasian perhitungan *hit ratio* sebesar 84.71%, serta hasil perhitungan berdasarkan variabel signifikan bahwa pengklasifikasian perusahaan pada kelompok *financial distress* atau kelompok bangkrut dan kelompok *non financial distress* atau kelompok tidak bangkrut menggunakan fungsi regresi logistik biner dalam penelitian ini

adalah stabil atau akurat, karena nilai *Press's Q* lebih besar dari ( $\chi^2_{tabel} = 81.91$ ) > ( $\chi^2_{(1,0.05)} = 3.84146$ ). Sedangkan rata-rata ketepatan prediksi AD model 2 sebesar 86.95% peluang kesalahan klasifikasi sebesar 0.0706, presentase ketepatan pengklasifikasian perhitungan *hit ratio* sebesar 92.94%, serta hasil perhitungan berdasarkan variabel signifikan dan variabel yang tidak signifikan dimasukkan kedalam model bahwa pengklasifikasian perusahaan pada kelompok *financial distress* atau kelompok bangkrut dan kelompok *non financial distress* atau kelompok tidak bangkrut menggunakan fungsi regresi logistik biner dalam penelitian ini adalah stabil atau akurat, karena nilai *Press's Q* lebih besar dari ( $\chi^2_{tabel} = 125.39$ ) > ( $\chi^2_{(1,0.05)} = 3.84146$ ).

### 3.4 Regresi Logistik Biner

Berdasarkan uji hipotesis untuk uji parsial didapatkan bahwa terdapat delapan parameter  $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7$  dan  $\beta_8$  yang tidak signifikan. Sehingga akan dilakukan perhitungan regresi ulang dengan mengeluarkan parameter yang tidak signifikan atau akan dilakukan perhitungan regresi logistik biner ulang dengan menggunakan parameter yang signifikan yaitu parameter  $\beta_9$ . Sehingga dapat disimpulkan ada satu model regresi logistik biner yang terbentuk dengan variabel signifikan yaitu variabel NPM ( $X_9$ ):

$$\hat{\pi}_i = \frac{\exp(0.541 + 0.107X_{9i})}{1 + \exp(0.541 + 0.107X_{9i})}$$

Berikut ini merupakan model regresi logistik biner yang terbentuk dengan variabel yang signifikan dan variabel yang tidak signifikan, sebagai berikut :

$$\hat{\pi}_i = \frac{\exp\left(\begin{array}{l} 5.586 - 0.739X_{1i} + 0.123X_{2i} - 61.164X_{3i} + 31.547X_{4i} - 0.539X_{5i} + 0.248X_{6i} \\ + 0.042X_{7i} - 0.030X_{8i} - 0.153X_{9i} \end{array}\right)}{1 + \exp\left(\begin{array}{l} 5.586 - 0.739X_{1i} + 0.123X_{2i} - 61.164X_{3i} + 31.547X_{4i} - 0.539X_{5i} + 0.248X_{6i} \\ + 0.042X_{7i} - 0.030X_{8i} - 0.153X_{9i} \end{array}\right)}$$

Model regresi logistik biner menunjukkan bahwa rata-rata ketepatan prediksi RLB model 1 sebesar 81.3%, peluang kesalahan klasifikasi sebesar 0.0647, presentase ketepatan pengklasifikasian perhitungan *hit ratio* sebesar 93.53%, serta hasil perhitungan berdasarkan variabel signifikan bahwa pengklasifikasian perusahaan pada kelompok *financial distress* atau kelompok bangkrut dan kelompok *non financial distress* atau kelompok tidak bangkrut menggunakan fungsi regresi logistik biner dalam penelitian ini adalah stabil atau akurat, karena nilai *Press's Q* lebih besar dari ( $\chi^2_{tabel} = 128.85$ ) > ( $\chi^2_{(1,0.05)} = 3.84146$ ). Sedangkan rata-rata ketepatan prediksi RLB model 2 sebesar 88.3%, peluang kesalahan klasifikasi sebesar 0.0588, presentase ketepatan pengklasifikasian perhitungan *hit ratio* sebesar 94.12%, serta hasil perhitungan

berdasarkan variabel signifikan dan variabel yang tidak signifikan dimasukkan kedalam model bahwa pengklasifikasian perusahaan pada kelompok *financial distress* atau kelompok bangkrut dan kelompok *non financial distress* atau kelompok tidak bangkrut menggunakan fungsi regresi logistik biner dalam penelitian ini adalah stabil atau akurat, karena nilai *Press's Q* lebih besar dari ( $\chi^2_{tabel} = 132.35$ ) > ( $\chi^2_{(1,0.05)} = 3.84146$ ).

#### 4 KESIMPULAN

Berdasarkan kedua uji yaitu analisis diskriminan dan logistik biner dapat disimpulkan bahwa hasil kedua uji tersebut tidak dapat dibandingkan. Hasil uji analisis diskriminan terpilih dua variabel yang signifikan berpengaruh terhadap tingkat potensi kebangkrutan perusahaan properti dan *real estate* pada taraf pengujian 5%. yaitu  $\ln DER$  dan  $\sqrt{GPM}$  sedangkan hasil uji regresi logistik biner menunjukkan bahwa hanya ada satu variabel yaitu NPM ( $X_9$ ) berpengaruh terhadap tingkat potensi kebangkrutan pada taraf pengujian 5%. Oleh karena variabel signifikan yang dihasilkan dari kedua analisis tersebut berbeda maka hasil uji kedua analisis tidak dapat dibandingkan.

Hasil perbandingan berdasarkan variabel yang signifikan dan variabel yang tidak signifikan dimasukkan ke dalam model analisis regresi logistik biner dan model analisis diskriminan.

**Tabel 6 Hasil Perbandingan Analisis Diskriminan dan Regresi Logistik Biner**

Keterangan	Analisis Regresi Logistik Biner	Analisis Diskriminan
Rata-Rata Ketepatan Prediksi	88.3%	86.95%
APER	0.0588	0.0706
<i>Hit Ratio</i>	94.12%	92.94%
<i>Press's Q</i>	132.35	125.39

Berdasarkan kedua analisis tersebut dapat disimpulkan bahwa hasil analisis regresi logistik biner memiliki rata-rata ketepatan prediksi lebih besar dari rata-rata ketepatan prediksi analisis diskriminan.

#### 5 DAFTAR PUSTAKA

- [1] Agresti, A. 2007. *An Introduction To Categorical Data Analysis Second Edition*. Wiley-Interscience: A John Willey & Sons, Inc. Publication.
- [2] Azkia, F. 2016. *Indeks Harga Properti Triwulan III-2016 Diprediksi Naik*. Diakses tanggal 12 April 2017. <http://www.Liputan6.com>
- [3] Clardiesty, R. 2017. *Analisis Properti Di Kota Depok, Jawa Barat April 2013*. Diakses tanggal 25 Mei 2017. <https://blog.urbanindo.com/2013/05/ analisis-properti-di-kota-depok-jawa-barat-april-2013/>

- [4] Everitt, B. S. 2005. *An Rand S-PLUS Companion to Multivariate Analysis*, Springer, London.
- [5] Hosmer, D. W, S. Lemeshow and R. Studivant. 2013. *Applied Logistic Regression Third Edition*, John Willey & Son, Inc: New York.
- [6] Johnson, R. A. and D. W. Wichern. 2002. *Applied Multivariate Statistical Analysis*, Fifth Edition, Prentice-Hall, Inc. New Jersey.
- [7] Nuruwael, G.M. dan S. Sihotang. 2013. Analisis Rasio Keuangan Sebagai Alat Untuk Menilai Kinerja Keuangan PT. Interbational Nickel Corporation, Tbk. *Jurnal Ilmu dan Riset Manajemen*. Vol. 2. No. 1. Sekolah Tinggi Ilmu Ekonomi Indonesia (STIESIA). Surabaya. Diakses tanggal 12 April 2017. [repository.stiesia.ac.id/768/1/HALAMAN%20COVER.pdf](http://repository.stiesia.ac.id/768/1/HALAMAN%20COVER.pdf)

**Lampiran 1. Ketepatan Klasifikasi Analisis Diskriminan Berdasarkan Variabel Signifikan**

Y	Prediksi Keanggotaan Grup		Total
	Financial Distress atau Bangkrut	Non Financial Distress atau Tidak Bangkrut	
Financial Distress atau Bangkrut	20	5	25
Non Financial Distress atau Tidak Bangkrut	21	124	145
Financial Distress atau Bangkrut (%)	80.0	20.0	100.0
Non Financial Distress atau Tidak Bangkrut (%)	14.5	85.5	100.0

**Lampiran 2. Ketepatan Klasifikasi Analisis Diskriminan Berdasarkan Variabel Signifikan dan Variabel Tidak Signifikan**

Y	Prediksi Keanggotaan Grup		Total
	Financial Distress atau Bangkrut	Non Financial Distress atau Tidak Bangkrut	
Financial Distress atau Bangkrut	16	9	25
Non Financial Distress atau Tidak Bangkrut	3	142	145
Financial Distress atau Bangkrut (%)	64.0	36.0	100.0
Non Financial Distress atau Tidak Bangkrut (%)	2.1	97.9	100.0

**Lampiran 3. Ketepatan Klasifikasi Analisis Regresi Logistik Biner Berdasarkan Variabel Signifikan**

Y	Prediksi Keanggotaan Grup		Persentase yang benar (%)
	Financial Distress atau Bangkrut	Non Financial Distress atau Tidak Bangkrut	
Financial Distress atau Bangkrut	16	9	64.0
Non Financial Distress atau Tidak Bangkrut	2	143	98.6
Persentase keseluruhan (%)			93.5

**Lampiran 4. Ketepatan Klasifikasi Analisis Regresi Logistik Biner Berdasarkan Variabel Signifikan dan Variabel Tidak Signifikan**

Y	Prediksi Keanggotaan Grup		Persentase yang benar (%)
	Financial Distress atau Bangkrut	Non Financial Distress atau Tidak Bangkrut	
Financial Distress atau Bangkrut	20	5	80.0
Non Financial Distress atau Tidak Bangkrut	5	140	96.6
Persentase keseluruhan (%)			94.1

# Aplikasi Analisis *Survival* Dengan Regresi *Cox* Pada Pasien Penderita Penyakit Malaria

(Studi Kasus: Pasien Penderita Penyakit Malaria Di Rsu Dr M Yunus Tahun 2014)

Reti Engga Sari<sup>1</sup>, Sigit Nugroho<sup>2</sup>, dan Etis Sunandi<sup>2</sup>

<sup>1</sup>Alumni Jurusan Statistika FMIPA Universitas Bengkulu

<sup>2</sup>Staf Pengajar Jurusan Statistika FMIPA Universitas Bengkulu

Abstrak-Penelitian ini bertujuan untuk mengetahui faktor-faktor yang mempengaruhi laju kesembuhan pasien penderita penyakit malaria. Analisis yang digunakan untuk mengetahui faktor-faktor yang mempengaruhi laju kesembuhan pasien penderita penyakit malaria adalah analisis *survival* dengan pemodelan regresi *cox*. Data yang digunakan pada penelitian ini berupa data sekunder yang diperoleh dari data rekam medis di RSU Dr. M. Yunus Bengkulu pada Tahun 2014 dengan variabel respon berupa waktu *survival*. Waktu *survival* yang digunakan adalah waktu lamanya pasien dirawat. Distribusi untuk waktu *survival* yaitu Distribusi Log-Normal. Hasil analisis *survival* dengan pemodelan regresi *cox* menunjukkan bahwa terdapat lima faktor yang berpengaruh secara signifikan terhadap laju kesembuhan pasien penderita penyakit malaria yaitu umur, kadar hematokrit, jumlah trombosit, jumlah leukosit, dan malaria komplikasi.

Kata Kunci: malaria, analisis *survival*, regresi *cox*, Distribusi Log-Normal.

## I. PENDAHULUAN

Malaria merupakan salah satu penyakit infeksi yang dapat mengancam jiwa. Penyakit ini disebabkan oleh gigitan nyamuk *Anopheles species* betina yang ditularkan pada manusia baik pada anak-anak maupun orang dewasa. Malaria saat ini masih menjadi masalah kesehatan masyarakat, baik di Indonesia maupun di negara lain.

Menurut *The World Malaria Report* (2012), malaria masih menjadi masalah kesehatan utama di 106 negara. Malaria menyerang sedikitnya 350 – 500 juta penduduk setiap tahun dan menyebabkan kematian sekitar 1 juta jiwa pertahun. *The World Malaria Report* menyatakan bahwa “masih sekitar 3,2 miliar orang bertempat tinggal di daerah endemis malaria”.

Menurut Yatim (2007), penyakit malaria ditandai dengan demam yang sering berkala (periodik) dengan berbagai derajat, kurang darah (anemia), limpa membesar, serta dengan

berbagai kelompok gejala (sindroma) lain, seperti gangguan pada hati, otak, dan ginjal. Jika tidak diobati, malaria dapat dengan cepat mengancam jiwa karena mengganggu aliran darah ke organ vital.

Dalam kajian klinis, akan diamati respon terhadap pengobatan yang sudah diberikan kepada pasien. Seiring dengan waktu, pasien dapat dinyatakan sembuh jika pasien tersebut terbebas dari gejala-gejala penyakitnya. Sedangkan sebagian pasien lain dapat mengalami keadaan sakit yang berulang kembali bahkan meninggal dunia (Kudus, Muchlis, dan Respati, 2011). Peristiwa pertama kali pasien dinyatakan menderita penyakit malaria sampai waktu tertentu yang ingin diteliti (seperti kesembuhan, keadaan sakit yang berulang kembali, dan kematian) termasuk kedalam analisis *survival*.

Analisis *survival* merupakan suatu metode yang berkaitan dengan waktu, mulai dari *time origin* atau *start point* sampai dengan terjadinya

suatu kejadian khusus (*failure event* atau *end point*). Salah satu metode dari analisis *survival* yang sering digunakan dari model regresi yaitu regresi *cox*. Regresi *cox* merupakan salah satu analisis *survival* yang dapat digunakan untuk mengetahui hubungan diantara variabel dependen dengan variabel independen yang merupakan waktu *survival* (Collet, 2003).

Oleh karena itu, melalui penelitian ini akan dilakukan analisis *survival* dengan pemodelan regresi *cox* terhadap faktor-faktor yang diduga mempengaruhi laju kesembuhan pasien penderita penyakit malaria di RSUD Dr. M. Yunus Bengkulu. Pemilihan di RSUD Dr. M. Yunus sebagai tempat penelitian karena RSUD Dr. M. Yunus merupakan rumah sakit terbesar dan terlengkap di Provinsi Bengkulu.

## II. TINJAUAN PUSTAKA

### 2.1 Analisis Survival

Analisis *survival* merupakan suatu metode yang berkaitan dengan waktu, mulai dari *time origin* atau *start point* sampai dengan terjadinya suatu kejadian khusus (*failure event* atau *end point*). Analisis *survival* memerlukan data yang merupakan waktu *survival* (Collet, 2003).

Terdapat tiga hal yang harus diperhatikan dalam menentukan waktu *survival* ( $T$ ), yaitu (Kleinbaum and Klein, 2005):

1. *Time origin* atau *starting point* (titik awal) adalah waktu dimulainya penelitian.
2. *Ending event of interest* (kejadian akhir) adalah kejadian yang menjadi tujuan penelitian.
3. *Measurement scale for the passage of time* (skala ukuran untuk berlalunya waktu).

### 2.2 Regresi Cox

Regresi *cox* merupakan salah satu analisis regresi yang sering digunakan pada analisis *survival*. Respon yang digunakan adalah data yang diperoleh dari perhitungan waktu suatu peristiwa tertentu, misalnya kematian, kesembuhan, dan kekambuhan. Model dari regresi *cox* adalah sebagai berikut (Kleinbaum and Klein, 2005):

$$h(t, X) = h_o(t)e^{\sum_{i=1}^p \beta_i X_i} \\ = h_o(t)\exp(\beta_1 X_1 + \beta_2 X_2 \dots + \beta_p X_p) \quad (2.1)$$

dengan:

$h(t, X)$  = resiko kegagalan individu pada waktu  $t$  dengan karakteristik  $X$

$h_o(t)$  = fungsi *hazard* dasar ( fungsi *baseline hazard*)

$X_i$  = variabel penjelas

$\beta_i$  = koefisien parameter peubah penjelas  $X_i$  dalam model

$i$  = banyaknya parameter dalam model,  $i = 1, 2, \dots, p$

$t$  = waktu terjadinya *event*

### 2.3 Pendugaan Distribusi

Pengujian distribusi data pada analisis *survival* dapat dilakukan dengan menggunakan metode *Anderson-Darling*. Metode ini digunakan untuk menguji apakah data yang digunakan dalam analisis berasal dari distribusi tertentu, seperti Eksponensial, Log-Normal, *Weibull*, Log-Logistik dan lain-lain (Budiantara dan Shofa, 2012).

Uji *Anderson-Darling* dapat diperoleh dengan menggunakan rumus sebagai berikut:

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i - 1) \left[ \ln F(Y_i) + \ln \left( 1 - F(Y_{n+1-i}) \right) \right] \quad (2.2)$$

dengan:

$A^2$  = statistik uji untuk metode *Anderson-Darling*

$n$  = ukuran sampel  
 $x_i$  = data waktu *survival* yang telah diurutkan  
 $\bar{x}$  = rata-rata data waktu *survival*  
 $s$  = standar deviasi data waktu *survival*  
 $F(Y_i)$  = nilai fungsi distribusi kumulatif dari distribusi tertentu

## 2.4 Estimasi dan Pengujian Signifikansi Parameter

Uji parameter dilakukan untuk mengetahui signifikansi parameter terhadap respon atau model. Uji parameter ada dua yaitu:

- (1) Uji serentak, uji ini digunakan untuk mengetahui apakah variabel independen yang digunakan pada model berpengaruh signifikan secara bersama-sama. Berikut hipotesisnya (Wulandari, Salamah, dan Susilaningrum, 2009):

$$H_0: \beta_1 = \beta_2 = \dots = \beta_p$$

$$H_1: \text{paling sedikit ada satu } \beta_i \neq \beta_j, i \neq j = 1, 2, \dots, p$$

$$G^2 = 2(\ln l_1 - \ln l_0) = 2(L_1 - L_0) \quad (2.3)$$

dengan:

$L_1$  = nilai *log-likelihood* dengan variabel bebas

$L_0$  = nilai *log-likelihood* tanpa variabel bebas

Statistik uji  $G^2$  merupakan *likelihood ratio test*. Keputusan: tolak  $H_0$  jika nilai statistik uji  $G^2 > \chi^2_{p,\alpha}$  dengan  $p$  derajat bebas, dimana  $p$  merupakan banyaknya parameter dalam model tanpa  $\beta_0$ .

- (2) Uji parsial, uji ini digunakan untuk mengetahui variabel independen mana yang berpengaruh signifikan terhadap

model. Berikut hipotesisnya (Wulandari, Salamah, dan Susilaningrum, 2009):

$$H_0: \beta_k = 0$$

$$H_1: \beta_k \neq 0, \text{ dengan } k = 1, 2, \dots, n$$

$$W_k^2 = \left[ \frac{\widehat{\beta}_k}{SE(\widehat{\beta}_k)} \right]^2 \quad (2.4)$$

dengan:

$$W_k^2 = \text{statistik uji wald}$$

$$\widehat{\beta}_k = \text{penduga } \beta_k$$

$$SE(\widehat{\beta}_k) = \text{standar error dari } \beta_k$$

Keputusan: tolak  $H_0$  jika nilai statistik uji

$$W_k > \chi^2_{1,\alpha}.$$

## 2.5 Pemilihan Model Terbaik dan Odds Ratio

Pemilihan model terbaik dilakukan untuk membandingkan sejumlah kemungkinan model, hal ini dapat dilakukan dengan menghitung nilai *Akaike Information Criterion (AIC)*. Model terbaik merupakan model yang memiliki nilai *AIC* terkecil. (Collet, 2003). Nilai *AIC* dapat diperoleh dari:

$$AIC = -2 \log \hat{L} + 2k \quad (2.5)$$

dengan:

$$\hat{L} = \text{fungsi } \textit{likelihood}.$$

$k$  = jumlah parameter dalam model yang diduga.

Laju kesembuhan pasien dapat diketahui melalui nilai *odds ratio*. Rumus *odds ratio* adalah sebagai berikut (Kleinbaum and Klein, 2005):

$$OR = \frac{h_0(t|x=0)}{h_0(t|x=1)} = \frac{h_0(t)e^{(\beta)}}{h_0(t)} = e^{(\beta)} \quad (2.6)$$

## III. METODE PENELITIAN

### 3.1 Lokasi dan Waktu

Unit observasi yang digunakan dalam penelitian ini adalah banyaknya jumlah pasien penderita penyakit malaria di RSUD Dr. M.

Yunus Bengkulu. Waktu penelitian yang digunakan pada penelitian yaitu pasien penderita penyakit malaria di RSUD Dr. M. Yunus Bengkulu dari bulan Januari 2014 – Desember 2014.

### 3.2 Sumber dan Jenis Data

Sumber data pada penelitian ini berupa data sekunder yang merupakan data rekam medis dari RSUD Dr. M. Yunus Bengkulu mengenai waktu *survival* pasien penderita penyakit malaria pada tahun 2014. Waktu *survival* adalah lamanya pasien malaria dirawat di RSUD Dr. M. Yunus Bengkulu.

### 3.3 Variabel Dependen dan Independen

Variabel dependen yang digunakan dalam penelitian ini adalah waktu lamanya pasien penderita penyakit malaria dirawat, ( $Y$ ). Sedangkan variabel independen yang diduga mempengaruhi laju kesembuhan pasien penderita penyakit malaria adalah sebagai berikut: Umur ( $X_1$ ), Jenis kelamin ( $X_2$ ), Jumlah Hemoglobin ( $X_3$ ), Kadar Hematokrit ( $X_4$ ), Jumlah Trombosit ( $X_5$ ), Jumlah Leukosit ( $X_6$ ), dan Malaria komplikasi ( $X_7$ ).

### 3.4 Prosedur Penelitian

Prosedur yang dilakukan dalam penelitian ini adalah sebagai berikut:

1. Mengambil data di ruang rekam medis RSUD Dr. M. Yunus Bengkulu.
2. Deskriptif dan karakteristik data analisis *survival*.
3. Pengujian distribusi data waktu *survival* dengan menggunakan metode *Anderson-Darling*.
4. Menghitung Estimasi Parameter.
5. Pengujian Signifikansi Parameter.

6. Pemilihan model terbaik.
7. Pemodelan Regresi *Cox*.
8. Menghitung Nilai *Odds Ratio*
9. Menarik kesimpulan.

## IV. HASIL DAN PEMBAHASAN

### 4.1 Deskripsi dan Karakteristik Pasien Penderita Penyakit Malaria

Data *survival* dalam penelitian analisis *survival* pada pasien penderita penyakit malaria di RSUD Dr. M. Yunus tahun 2014 sebanyak 61 data.

#### 4.1.1 Deskripsi Waktu *Survival*

Berikut hasil yang menunjukkan deskripsi mengenai karakteristik waktu *survival* pasien penderita penyakit malaria.

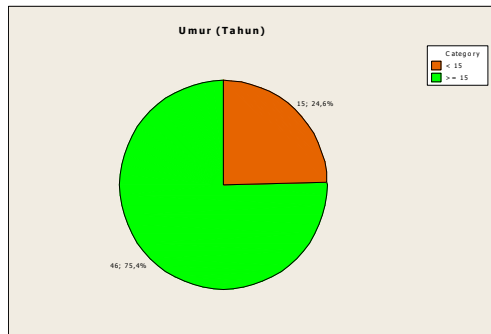
Tabel 1. Deskripsi Waktu *Survival*

Deskriptif Waktu <i>Survival</i>	Lama Pasien di Rawat (hari)
<i>Mean</i>	4.541
St. Deviasi	2.500
Minimum	2.000
Median	4.000
Maksimum	14.000

Tabel 1 menunjukkan bahwa waktu *survival* pada pasien penderita penyakit malaria memiliki rata-rata waktu 4 hari dengan standar deviasi 2.500, waktu minimum 2 hari, waktu maksimum 14 hari dan median waktu sebesar 4.

#### 4.1.2 Umur ( $X_1$ )

Persentase pasien penderita penyakit malaria berdasarkan umur dapat dilihat pada gambar 4.1.

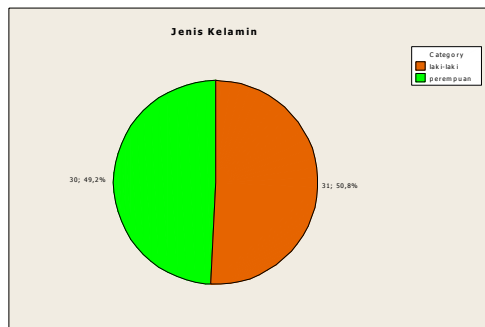


Gambar 4.1. Karakteristik Umur

Gambar 4.1. menunjukkan bahwa persentase pasien penderita penyakit malaria berumur 0-14 tahun sebesar 24.6 % (15 orang), sedangkan persentase pasien berumur 15-80 tahun sebesar 75.4 % (46 orang). Terlihat bahwa persentase pasien penderita penyakit malaria berumur 0-14 cenderung lebih rendah daripada pasien berumur 15-80 tahun.

#### 4.1.3 Jenis Kelamin ( $X_2$ )

.Persentase pasien penderita penyakit malaria berdasarkan jenis kelamin dapat dilihat pada gambar 4.2.

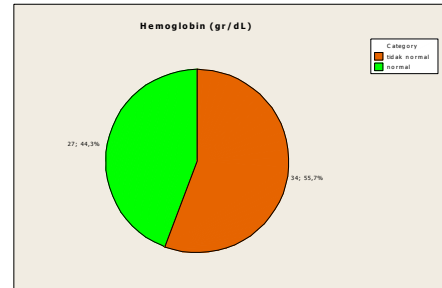


Gambar 4.2. Karakteristik Jenis Kelamin

Gambar 4.2. menunjukkan bahwa persentase pasien penderita penyakit malaria berjenis kelamin laki-laki sebesar 50.8 % (31 orang), sedangkan persentase pasien berjenis kelamin perempuan sebesar 49.2 % (30 orang). Terlihat bahwa persentase pasien penderita penyakit malaria berjenis kelamin laki-laki cenderung lebih tinggi daripada pasien perempuan.

#### 4.1.4 Jumlah Hemoglobin ( $X_3$ )

Persentase pasien penderita penyakit malaria berdasarkan jumlah hemoglobin dapat dilihat pada gambar 4.3.

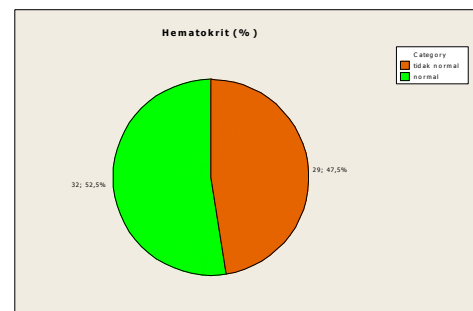


Gambar 4.3 Karakteristik Jumlah Hemoglobin

Gambar 4.3. menunjukkan bahwa persentase pasien penderita penyakit malaria dengan jumlah hemoglobin normal sebesar 44.3 % (27 orang), sedangkan pasien penderita penyakit malaria dengan jumlah hemoglobin tidak normal sebesar 55.7 % (34 orang). Terlihat bahwa persentase pasien penderita penyakit malaria dengan jumlah hemoglobin normal cenderung lebih rendah daripada pasien dengan jumlah hemoglobin tidak normal.

#### 4.1.5 Kadar Hematokrit ( $X_4$ )

Persentase pasien penderita penyakit malaria berdasarkan jumlah hematokrit dapat dilihat pada gambar 4.4.



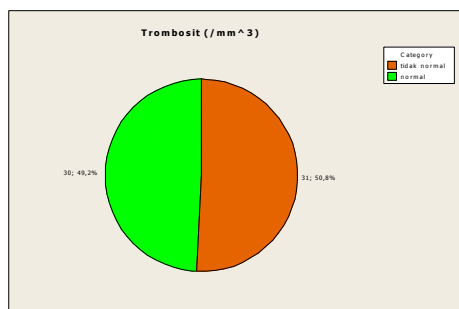
Gambar 4.4. Karakteristik Kadar Hematokrit

Gambar 4.4. menunjukkan bahwa persentase pasien penderita penyakit malaria dengan kadar hematokrit dalam jumlah normal sebesar 52.5 % (32 orang), sedangkan pasien penderita penyakit

malaria dengan kadar hematokrit dalam jumlah tidak normal sebesar 47.5 % dengan (29 orang). Terlihat bahwa persentase pasien penderita penyakit malaria dengan jumlah hematokrit normal cenderung lebih tinggi daripada pasien dengan jumlah hematokrit tidak normal.

#### 4.1.6 Jumlah Trombosit ( $X_5$ )

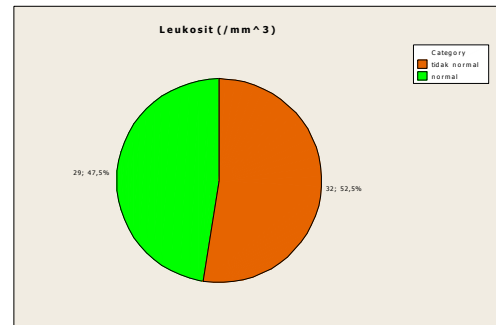
Persentase pasien penderita penyakit malaria berdasarkan jumlah trombosit dapat dilihat pada gambar 4.5.



Gambar 4.5. Karakteristik Jumlah Trombosit  
 Gambar 4.5. menunjukkan bahwa persentase pasien penderita penyakit malaria dengan jumlah trombosit normal sebesar 49.2 % (30 orang), sedangkan pasien penderita penyakit malaria dengan jumlah trombosit tidak normal sebesar 50.8 % (31 orang). Terlihat bahwa persentase pasien penderita penyakit malaria dengan jumlah trombosit normal cenderung sama dengan jumlah trombosit tidak normal.

#### 4.1.7 Jumlah Leukosit ( $X_6$ )

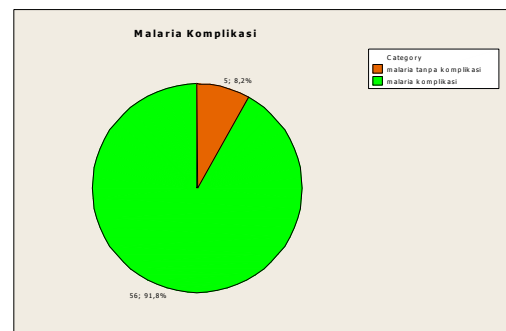
Persentase pasien penderita penyakit malaria berdasarkan jumlah leukosit dapat dilihat pada gambar 4.6.



Gambar 4.6. Karakteristik Jumlah Leukosit  
 Gambar 4.6. menunjukkan bahwa persentase pasien penderita penyakit malaria dengan jumlah leukosit normal sebesar 47.5 % (29 orang), sedangkan pasien penderita penyakit malaria dengan jumlah trombosit tidak normal sebesar 52.5 % (32 orang). Terlihat bahwa persentase pasien penderita penyakit malaria dengan jumlah leukosit normal cenderung lebih rendah daripada jumlah leukosit tidak normal.

#### 4.1.8 Malaria Komplikasi ( $X_7$ )

Persentase pasien penderita penyakit malaria berdasarkan jumlah leukosit dapat dilihat pada gambar 4.7.



Gambar 4.7. Karakteristik Malaria Komplikasi  
 Gambar 4.7. menunjukkan persentase jenis malaria yang diderita oleh pasien yang meliputi, pasien yang menderita malaria tanpa komplikasi dengan persentase sebesar 8.2 % (5 orang), sedangkan pasien yang menderita malaria komplikasi memiliki persentase jauh lebih besar yaitu 91.8 % (56 orang). Terlihat bahwa persentase pasien yang menderita malaria tanpa

komplikasi jauh lebih rendah daripada persentase pasien yang menderita malaria komplikasi.

#### 4.2 Pengujian Distribusi Waktu *Survival*

Berdasarkan pedugaan distribusi dengan menggunakan uji *Anderson-Darling*, diketahui bahwa nilai *Anderson-Darling* terkecil terdapat pada Distibusi Log-Normal. Berikut hasil pengujian yang diperoleh dengan menggunakan *software minitab 16*:

Tabel 2. Pengujian Distribusi Waktu *Survival*

Distribusi	Nilai <i>Anderson-Darling</i>
<i>Exponential</i>	7.842
Log-Normal	1.645
<i>Weibull</i>	2.043
Log-Logistik	1.715

Fungsi *hazard* yang digunakan sebagai *baseline hazard* dalam pemodelan regresi *Cox* dengan pendekatan Distribusi Log-Normal adalah sebagai berikut:

$$h(t) = \frac{1}{t\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln t - \mu)^2}{2\sigma^2}\right) \frac{1}{1 - \Phi\left(\frac{\ln t - \mu}{\sigma}\right)}$$

Tabel berikut merupakan rangkuman hasil estimasi waktu *survival* pasien penderita penyakit malaria dengan pendekatan Distribusi Log-Normal:

Tabel 3. Estimasi Parameter Waktu *Survival*

Parameter	$\widehat{\beta}_k$	$SE(\widehat{\beta}_k)$
<i>Location</i>	1.384	0.064
<i>Scale</i>	0.498	0.045

*Log-Likelihood* = -128.491

Berdasarkan hasil estimasi waktu *survival* pada Tabel 3, diperoleh hasil estimasi dari  $\mu$  adalah 1.384 dan estimasi dari parameter  $\sigma$  adalah

0.498 sehingga persamaan fungsi *hazard* adalah sebagai berikut:

$$h(t) = \frac{1}{0.498t\sqrt{2\pi}} \exp\left(-\frac{(\ln t - 1.384)^2}{1.798}\right) \frac{1}{1 - \Phi\left(\frac{\ln t - 1.384}{0.498}\right)}$$

#### 4.3 Estimasi Parameter dan Pengujian Signifikasi Parameter

Tabel 4. Estimasi Parameter dan Pengujian Signifikasi Parameter

Variabel	Db	$\widehat{\beta}_k$	$SE(\widehat{\beta}_k)$	$W_k^2$	Signifikasi	Keputusan
$(X_1)$	1	-0.329	0.137	5.757	0.016	Tolak $H_0$
$(X_2)$	1	-0.073	0.105	0.484	0.487	Terima $H_0$
$(X_3)$	1	-0.116	0.108	1.169	0.280	Terima $H_0$
$(X_4)$	1	-0.267	0.102	6.842	0.009	Tolak $H_0$
$(X_5)$	1	-0.260	0.104	6.237	0.013	Tolak $H_0$
$(X_6)$	1	-0.223	0.102	4.792	0.029	Tolak $H_0$
$(X_7)$	1	0.732	0.233	9.845	0.002	Tolak $H_0$

*Log-Likelihood* = -112.486

##### 1. Uji serentak

Hipotesis:

$$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7$$

$$H_1: \text{paling sedikit ada satu } \beta_i \neq \beta_j$$

dengan:

$$i \neq j = 1, 2, 3, 4, 5, 6, 7$$

Taraf Signifikasi:  $\alpha = 0.05$

Statistik Uji :

$$G^2 = 32.010$$

Keputusan: Pada hasil analisis diperoleh nilai  $G^2$  sebesar 32.010, dengan tingkat signifikan sebesar 5% diperoleh  $\chi_{0.05,7}^2 = 14.067$ . Karena nilai  $G^2 = 32.010 > \chi_{0.05,7}^2 = 14.067$  maka  $H_0$  ditolak.

Kesimpulan:  $H_0$  ditolak artinya pada taraf signifikasi 5% paling sedikit ada satu  $\beta_i \neq \beta_j, i \neq j = 1, 2, 3, 4, 5, 6, 7$  yang berpengaruh terhadap model.

##### 2. Uji parsial

Hipotesis:

$$H_0: \beta_k = 0$$

$H_1: \beta_k \neq 0$ , dengan  $k = 1, 2, \dots, 7$

Taraf Signifikansi:  $\alpha = 0,05$

Statistik Uji: Tolak  $H_0$  jika nilai statistik uji  $W_k > x_{1,\alpha}^2$  atau  $Sig < \alpha$ .

Keputusan: Dengan tingkat signifikansi sebesar 5% diperoleh  $x_{1,0.05}^2$  sebesar 3.841. Berdasarkan hasil analisis pada Tabel 4., diperoleh bahwa  $(X_1)$ ,  $(X_4)$ ,  $(X_5)$ ,  $(X_6)$ , dan  $(X_7)$  signifikansi terhadap model.

Kesimpulan: Berdasarkan uji parsial dapat disimpulkan bahwa variabel independen yang berpengaruh signifikan terhadap model yaitu umur  $(X_1)$ , kadar hematokrit  $(X_4)$ , jumlah trombosit  $(X_5)$ , jumlah leukosit  $(X_6)$  dan malaria komplikasi  $(X_7)$ .

Tabel 5. Estimasi Parameter dan Pengujian Signifikansi Parameter

Variabel	Db	$\tilde{\beta}_k$	$SE(\tilde{\beta}_k)$	$W_k^2$	Signifikansi	Keputusan
$(X_1)$	1	-0.353	0.136	6.760	0.009	Tolak $H_0$
$(X_4)$	1	-0.276	0.101	7.429	0.006	Tolak $H_0$
$(X_5)$	1	-0.256	0.105	5.896	0.015	Tolak $H_0$
$(X_6)$	1	-0.204	0.100	4.138	0.042	Tolak $H_0$
$(X_7)$	1	0.762	0.219	12.090	0.001	Tolak $H_0$

$$\text{Log-Likelihood} = -113.219$$

1. Uji serentak, uji ini digunakan untuk mengetahui apakah variabel independen yang digunakan pada model berpengaruh signifikansi secara bersama-sama.

Hipotesis:

$$H_0: \beta_1 = \beta_4 = \beta_5 = \beta_6 = \beta_7$$

$H_1$ : paling sedikit ada satu  $\beta_i \neq \beta_j$

dengan:  $i \neq j = 1, 4, 5, 6, 7$

Taraf Signifikansi :  $\alpha = 0.05$

Statistik Uji :

$$G^2 = 30.544$$

Tolak  $H_0$  jika nilai statistik uji  $G^2 > x_{\alpha,p}^2$ .

Keputusan: Pada hasil analisis diperoleh nilai  $G^2$  sebesar 30.544, dengan tingkat signifikan sebesar 5% diperoleh  $\chi_{0.05,5}^2 = 11.070$ . Karena nilai  $G^2 = 30.544 > \chi_{0.05,5}^2 = 11.070$  maka  $H_0$  ditolak.

Kesimpulan:  $H_0$  ditolak artinya pada taraf signifikansi 5% paling sedikit ada satu  $\beta_i \neq \beta_j$ ,  $i \neq j = 1, 4, 5, 6, 7$  yang berpengaruh terhadap model.

2. Uji parsial, Hipotesis :  $H_0: \beta_k = 0$   
 $H_1: \beta_k \neq 0$ , dengan  $k = 1, 4, 5, 6, 7$

Taraf Signifikansi:  $\alpha = 0,05$

Statistik uji : Tolak  $H_0$  jika nilai statistik uji  $W_k > x_{1,\alpha}^2$  atau  $Sig < \alpha$ .

Keputusan: dengan tingkat signifikansi sebesar 5% diperoleh  $x_{1,0.05}^2$  sebesar 3.841. Berdasarkan hasil analisis pada Tabel 5, diperoleh bahwa  $(X_1)$ ,  $(X_4)$ ,  $(X_5)$ ,  $(X_6)$ , dan  $(X_7)$  signifikansi terhadap model.

Kesimpulan: berdasarkan uji parsial dapat disimpulkan bahwa variabel independen yang berpengaruh signifikan terhadap model yaitu umur  $(X_1)$ , kadar hematokrit  $(X_4)$ , jumlah trombosit  $(X_5)$ , jumlah leukosit  $(X_6)$  dan malaria komplikasi  $(X_7)$ .

#### 4.4 Pemilihan Model Terbaik

Berikut nilai  $AIC$  yang diperoleh:

Tabel 6. Nilai Akaike Information Criterion (AIC)

Variabel	Nilai Log-Likelihood	Nilai AIC
Semua Variabel	-112.486	238,972
$(X_1), (X_4), (X_5), (X_6),$ dan $(X_7)$	-113.219	236,438

Tabel 6 menunjukkan bahwa model terbaik adalah model dengan lima variabel, yaitu umur  $(X_1)$ , kadar hematokrit  $(X_4)$ , jumlah trombosit  $(X_5)$ , jumlah leukosit  $(X_6)$ , dan

malaria komplikasi ( $X_7$ ) dengan nilai AIC sebesar 236,438.

#### 4.5 Pemodelan Regresi Cox

Bentuk pemodelan regresi *cox* adalah sebagai berikut:

$$h(t, X) = \frac{1}{0.498t\sqrt{2\pi}} \exp\left(-\frac{(\ln t - 1.384)^2}{1.798}\right) \exp(-0.353X_1 - 0.276X_4 - 0.256X_5 - 0.204X_6 + 0.762X_7)$$

#### 4.6 Menghitung Nilai Odds Ratio

Berikut hasil *odds ratio* yang diperoleh:

Tabel 7. Nilai *Odds Ratio* untuk Variabel-Variabel yang Signifikan

Variabel	Db	$\hat{\beta}_k$	<i>Odds Ratio</i>
Umur ( $X_1$ )	1	-0.353	0.703
Kadar Hematokrit ( $X_4$ )	1	-0.276	0.759
Jumlah Trombosit ( $X_5$ )	1	-0.256	0.774
Jumlah Leukosit ( $X_6$ )	1	-0.204	0.815
Malaria Komplikasi ( $X_7$ )	1	0.762	2.143

## V. KESIMPULAN DAN SARAN

### 5.1 Kesimpulan

Berdasarkan hasil dan pembahasan yang telah diuraikan maka diperoleh kesimpulan sebagai berikut:

1. Variabel yang diduga mempengaruhi laju kesembuhan pasien penderita penyakit malaria menggunakan analisis *survival* dengan regresi *cox* adalah umur, kadar hematokrit, jumlah trombosit, jumlah leukosit dan malaria komplikasi.
2. Bentuk model regresi *cox* untuk pasien penderita penyakit malaria adalah sebagai berikut:

$$h(t, X) = \frac{1}{0.498t\sqrt{2\pi}} \exp\left(-\frac{(\ln t - 1.384)^2}{1.798}\right) \exp(-0.353X_1 - 0.276X_4 - 0.256X_5 - 0.204X_6 + 0.762X_7)$$

### 5.2 Saran

Berdasarkan analisis yang telah dilakukan, adapun saran yang dapat diberikan pada penelitian ini adalah:

1. Adanya penelusuran terhadap pemberian obat yang diberikan kepada pasien agar dapat ditambahkan sebagai faktor yang mempengaruhi laju kesembuhan pasien penderita penyakit malaria.
2. Untuk perkembangan lebih lanjut mengenai analisis *survival* dapat dilakukan perbandingan antara beberapa model dari analisis *survival*.

## DAFTAR PUSTAKA

- (1) Budiantara, I., dan N. F. Shofa. 2012. Analisis Survival dengan Pendekatan Multivariate Adaptive Regression Splines pada Kasus Demam Berdarah Dengue (DBD). Program Studi Statistika FMIPA ITS. *Jurnal Sains dan Seni ITS* Vol. 1, No. 1.
- (2) Collet, D. 2003. *Modelling Survival Data in Medical Research (Second Edition)*. Chapman & Hall. London.
- (3) Kleinbaum, D. G., and M. Klein. 2005. *Survival Analysis A Self-Learning Text (Second Edition)*. Springer. United States of America.
- (4) Kudus, A., R. D. Muchlis., dan T. Respati. 2011. Penaksiran Peluang Kesembuhan dengan Kekambuhan Berdistribusi Eksponensial. *Prosiding SNaPP2011 Sains, Teknologi, dan Kesehatan*, Vol. 2, No. 1.
- (5) World Health Organization (WHO), 2012. World Malaria Report. Switzerland: WHO; 8-9.
- (6) Wulandari, S. P., M. Salamah., dan D. Susilaningrum. 2009. *Diktat Pengajaran Analisis Data Kualitatif*. Institut Teknologi Sepuluh Nopember.
- (7) Yatim, F. 2007. *Macam-Macam Penyakit Menular dan Cara Pencegahannya*. Jilid 2. Pustaka Obor Populer. Jakarta.

# APLIKASI FUNGSI TRANSFER PADA DATA DEBIT AIR DI PLTA TES (Studi Kasus: PLTA Tes di Kabupaten Lebong Provinsi Bengkulu)

Melda Juliza<sup>1</sup>, Sigit Nugroho<sup>2</sup>, dan Jose Rizal<sup>3</sup>  
Program Studi Matematika  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Universitas Bengkulu  
Email: mhezawilantara@gmail.com

## ABSTRAK

Penelitian ini bertujuan untuk menentukan model dari data debit air terhadap data curah hujan dengan mengaplikasikan fungsi transfer berdasarkan data debit air dari tanggal 01 Januari sampai dengan 31 Desember 2013 dan meramalkan debit air sebanyak 3 periode ke depan. Penelitian diawali dengan pemodelan *Autoregressive Integrated Moving Average* (ARIMA) dan dilanjutkan dengan pemodelan fungsi transfer. Debit air sebagai deret output dan curah hujan sebagai deret input. Pemodelan ARIMA deret input diperoleh model terbaik berdasarkan kriteria MSE yaitu model ARIMA(0,1,2). Pemodelan fungsi transfer dilakukan dengan pemutihan deret input dan output, korelasi silang dan pembobotan respon impuls dan diperoleh model fungsi transfer terbaik FT(1,1,1)(1,1). Model fungsi transfer untuk data debit air adalah:

$$y_t = (1,53932)y_{t-1} + (0,59092)y_{t-2} + (0,23998)x_{t-1} - (0,27998)x_{t-2} + (0,07639)x_{t-3} + a_t - (0,76598)a_{t-1} + (0,77401)a_{t-2}$$

Hasil peramalan menggunakan model fungsi transfer (1,1,1)(1,1) yaitu tanggal 01 Januari 2014 sebesar  $39.07, 02 \frac{m^3}{s}$  Januari 2014 sebesar  $38.89 \frac{m^3}{s}$  dan 03 Januari 2014 sebesar  $38.41 \frac{m^3}{s}$ .

**Kata kunci:** debit air, curah hujan, ARIMA, fungsi transfer.

## I. PENDAHULUAN

PLTA adalah salah satu pembangkit yang memanfaatkan aliran air untuk diubah menjadi energi listrik. Pembangkit listrik ini bekerja dengan cara merubah energi air yang mengalir dari bendungan atau air terjun menjadi energi mekanik dengan bantuan turbin air dan dari energi mekanik menjadi energi listrik dengan bantuan generator, kemudian energi listrik tersebut dialirkan melalui jaringan-jaringan transmisi.

Danau Tes merupakan sumber air untuk pembangkit turbin PLTA Tes. Saat ini proses sedimentasi ke badan danau cukup mengkhawatirkan, dan apabila dilakukan pengerukan akan mengancam keberlangsungan PLTA Tes. Terjadinya pendangkalan oleh proses sedimentasi yang terus terjadi setiap waktu dapat mengancam fungsi Danau Tes sebagai sumber tenaga listrik (Haryanto, 2011).

Besarnya listrik yang dihasilkan oleh Pembangkit Listrik Tenaga Air (PLTA) tergantung pada dua hal, yaitu head (jarak tinggi air) dan debit (besar jumlah air yang mengalir). Semakin tinggi bendungan semakin tinggi air jatuh maka semakin besar tenaga yang dihasilkan, dan semakin banyak air yang jatuh maka turbin akan menghasilkan tenaga yang lebih banyak.

Pada operasi PLTA perhitungan keadaan air yang masuk pada penampungan air dan jumlah air yang tersedia didalamnya serta perhitungan besar air yang

akan dialirkan melalui pintu saluran air untuk menggerakkan turbin merupakan suatu keharusan untuk dimiliki. Dengan demikian kontrol terhadap air yang masuk maupun yang didistribusikan ke pintu saluran air untuk menggerakkan turbin harus dilakukan dengan baik, sehingga dalam operasi PLTA dapat dijadikan sebagai dasar tindakan pengaturan efisiensi penggunaan air maupun pengamanan seluruh sistem yang diharapkan dapat beroperasi sepanjang tahun, walaupun pada musim kemarau panjang. Jumlah air yang tersedia tergantung pada jumlah air yang mengalir di sungai.

Besar kecilnya debit air dipengaruhi oleh berbagai faktor, salah satunya dipengaruhi oleh curah hujan. Jika curah hujan di suatu daerah aliran sungai sedikit maka jumlah debit airnya akan berkurang (Wahid, 2009). Diketuainya hubungan antara curah hujan pada periode tertentu dengan debit air dapat dilakukan suatu peramalan. Fungsi transfer diharapkan dapat menjelaskan pengaruh curah hujan terhadap debit air.

Fungsi transfer adalah salah satu alternatif untuk menyelesaikan permasalahan bila terdapat lebih dari satu variabel deret berkala dimana salah satu variabel berpengaruh terhadap yang lain keadaan (Bowerman and Tukey, 1979), sedangkan model fungsi transfer merupakan model peramalan deret waktu yang menggabungkan karakteristik model ARIMA dengan pendekatan kausal (Makridakis dkk, 1999).

## II. TINJAUAN PUSTAKA

### A. Debit Air

Aliran sungai atau debit adalah jumlah air yang mengalir melalui suatu penampang sungai tertentu per satuan waktu. Satuan besaran debit dalam sistem satuan SI dinyatakan dalam satuan meter kubik per detik ( $m^3/det$ ). Besarnya debit air dapat diketahui dengan pengamatan dalam jangka waktu yang lama (Baros, 2009).

Debit air diukur dengan menggunakan *current-meter*, yaitu kecepatan aliran dan luas penampang basah diukur langsung di lapangan. Kecepatan aliran dihitung dengan mengukur jumlah putaran alat tersebut dalam suatu satuan waktu (Mursanto, 2010)

### B. Analisis Deret Waktu

Analisis deret waktu adalah analisis dari serangkaian data pengamatan yang terjadi berdasarkan indeks waktu secara berurutan dengan interval waktu tetap. Waktu atau periode yang dibutuhkan untuk melakukan suatu peramalan ini disebut sebagai *lead time* yang bervariasi pada tiap persoalan (Indrawati dan Sutijo, 2012).

Ciri-ciri dalam pembentukan model deret waktu adalah data stasioner. Apabila data tidak stasioner maka dilakukan *differencing* (pembedaan), yaitu deret asli diganti dengan deret selisih dengan persamaan:

$$\nabla^d X_t = (1 - B)^d X_t \quad (1)$$

Prosedur ARIMA meliputi empat tahapan untuk membentuk model ARIMA, yaitu identifikasi model dengan mengamati pola ACF dan PACF, penaksiran parameter, uji diagnosa model dan peramalan. Secara umum model ARIMA (p,d,q) dapat ditulis dalam bentuk

$$\phi_p(B)(1 - B)^d X_t = \theta_q(B) a_t \quad (2)$$

### C. Model Fungsi Transfer

Model fungsi transfer merupakan pengembangan dari model ARIMA. Jika deret berkala  $Y_t$  berhubungan dengan satu atau lebih deret berkala lain  $X_t$  maka dapat dibuat suatu model, model yang dihasilkan disebut fungsi transfer. Fungsi transfer adalah salah satu alternatif untuk menyelesaikan permasalahan jika terdapat lebih dari satu variabel berpengaruh terhadap yang lain keadaan. Bentuk umum model fungsi transfer *single input* adalah sebagai berikut (Makridakis dkk, 1999):

$$y_t = \frac{\omega_s(B)}{\delta_r(B)} x_{t-b} + n_t \quad (3)$$

Tahap-tahap pembentukan model fungsi transfer adalah:

#### 1. Prewhitening deret input

$$\alpha_t = \frac{\phi_x(B)}{\theta_x(B)} x_t \quad (4)$$

Dimana nilai  $\alpha_t$  adalah deret input yang telah mengalami prewhitening

#### 2. Prewhitening deret output

$$\beta_t = \frac{\phi_y(B)}{\theta_y(B)} y_t \quad (5)$$

#### 3. Fungsi korelasi silang (*cross correlation function*)

#### 4. Penetapan ( $b, r, s$ ) untuk model fungsi transfer

#### 5. Identifikasi Model deret noise ( $n_t$ )

Taksiran awal dari deret noise adalah:

$$n_t = y_t - v_0 x_t - v_1 x_{t-1} - v_2 x_{t-2} - \dots - v_g x_{t-g} \quad (6)$$

model sementara dari deret noise di atas dapat diidentifikasi dengan menyelidiki ACF dan PACF:

$$\phi_n(B) n_t = \theta_n(B) a_t \quad (7)$$

sehingga dengan mengkombinasikan kedua persamaan tersebut dapat diperoleh model fungsi transfer:

$$y_t = \frac{\omega_s(B)}{\delta_r(B)} x_{t-b} + \frac{\theta(B)}{\phi(B)} a_t \quad (8)$$

## III. METODE PENELITIAN

Data yang digunakan pada penelitian ini merupakan data sekunder yang diperoleh dari PLTA Tes dan BMKG Provinsi Bengkulu. Data tersebut adalah data harian debit air PLTA dan data harian curah hujan Kabupaten Lebong. Data yang digunakan berjumlah 365 data yaitu dari Bulan Januari sampai dengan Bulan Desember tahun 2013.

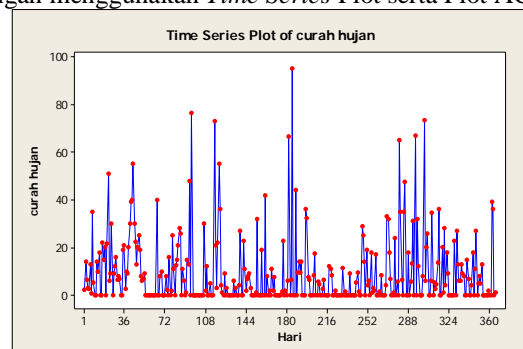
Metode analisis data yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Plot data untuk melihat kestasioneran data.
2. Identifikasi model fungsi transfer
  - a. Pengujian kestasioneran data
  - b. Pemutihan deret input dan deret output yang digunakan untuk memperoleh  $\alpha_t$  dan  $\beta_t$
  - c. Perhitungan korelasi silang dan pendugaan langsung bobot respons impuls yang digunakan untuk menentukan orde b,r,s.
  - d. Identifikasi konstanta (b,r,s) untuk model fungsi transfer yang menghubungkan deret input dan output
  - e. Penaksiran deret gangguan ( $n_t$ ) dan identifikasi model ARIMA ( $p_t, 0, q_t$ )
3. Penaksiran parameter-parameter model fungsi transfer
4. Uji diagnosa model fungsi transfer
  - a. Perhitungan korelasi silang nilai sisa model (b,r,s) dengan deret gangguan yang telah diputihkan
  - b. Perhitungan autokorelasi dari nilai sisa model (b,r,s)
5. Peramalan dengan model fungsi transfer

## IV. HASIL DAN PEMBAHASAN

### A. Pemodelan ARIMA Deret Input

Proses pemodelan ARIMA dapat dilakukan dengan menggunakan *Time Series Plot* serta *Plot ACF*



Gambar 1. Plot data asli curah hujan

Gambar 1. menunjukkan bahwa pada deret input curah hujan belum stasioner terhadap mean, sehingga untuk menstasionerkan data, maka perlu dilakukan *differencing*.

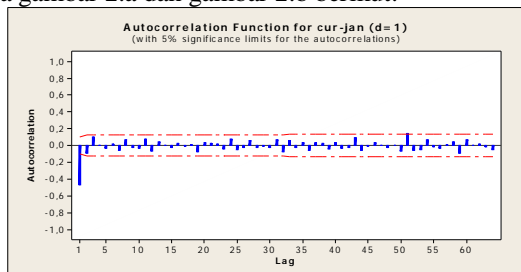
Setelah data di *differencing* maka dilakukan uji kestasioneran data dengan uji *Aughmented Dickey Fuller* (ADF), yaitu:

Tabel 1 Uji ADF debit air

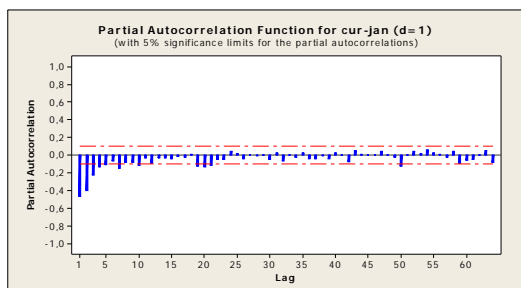
Variabel	nilai ADF	Nilai kritis MacKinnon (5%)
Debit air ( $y_t$ )	-15,39883	-3,422321
Curah hujan ( $x_t$ )	-15,19469	-3,422356

Berdasarkan Tabel 1. nilai ADF debit air adalah -15.39883 dan nilai ADF curah hujan adalah -15.19469, keduanya lebih kecil daripada nilai kritis statistik McKinon pada selang kepercayaan 5%. Artinya data debit air dan curah hujan telah stasioner pada *differencing* pertama.

Identifikasi model ARIMA dapat dilihat dari grafik ACF dan PACF dari data yang telah stasioner, seperti pada gambar 2.a dan gambar 2.b berikut:



Gambar 2.a Plot ACF deret  $x_t$



Gambar 2.b Plot PACF deret  $x_t$

Berdasarkan plot ACF terlihat bahwa grafik terpotong pada *lag* pertama namun pada *lag* kedua tidak menurun drastis, grafik PACF menurun secara bertahap maka diestimasi model MA(1) atau MA(2). Berikut hasil uji parameter untuk memperoleh model ARIMA terbaik:

Tabel 2. Pendugaan parameter ARIMA curah hujan

ARIMA	Parameter	T hitung	P-Value	MSE
(0,1,2)	MA 1 0,9305	156871,12	0,000	215,2
	MA 2 0,0590	5,59	0,000	
(1,1,2)	AR 1 -0,5724	-12,73	0,000	215,7
	MA 1 0,3437	20,85	0,000	

	MA 2 0,6428	75,47	0,000	
--	-------------	-------	-------	--

Berdasarkan Tabel 2. diperoleh nilai MSE terkecil 215,2 yaitu pada model ARIMA (0,1,2) yang menunjukkan bahwa model relatif sudah dalam bentuk yang paling sederhana (*parsimonius*). Jadi, model terbaik pada data curah hujan adalah model ARIMA (0,1,2). Sehingga diperoleh persamaan model ARIMA pada curah hujan yaitu:

$$x_t = (1 - 0,9305B - 0,0590B^2)\alpha_t$$

#### B. Pemutihan Deret Input dan Output

Tahap pemutihan dilakukan berdasarkan model ARIMA deret input. Tahap ini digunakan unsur *white noise* model tersebut, diperoleh model pemutihan deret input  $x_t$ :

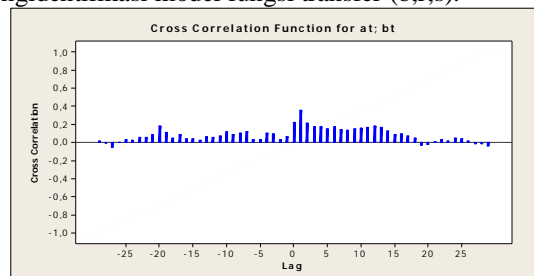
$$\alpha_t = x_t + 0,9305\alpha_{t-1} + 0,0590\alpha_{t-2}$$

Pemutihan deret output  $y_t$  diperoleh dengan cara melakukan transformasi yang sama dengan deret input, sehingga model pemutihan untuk deret output adalah:

$$\beta_t = y_t + 0,9305\beta_{t-1} + 0,0590\beta_{t-2}$$

#### C. Perhitungan Korelasi Silang dan Orde (b,r,s)

Tahap ini adalah tahap pembentukan awal pembentukan model fungsi transfer yaitu dengan pembuatan korelasi silang antara deret input dan output yang telah diputihkan. Korelasi silang menunjukkan hubungan antara curah hujan dengan debit air. Pola korelasi silang yang dihasilkan akan digunakan untuk mengidentifikasi model fungsi transfer (b,r,s).

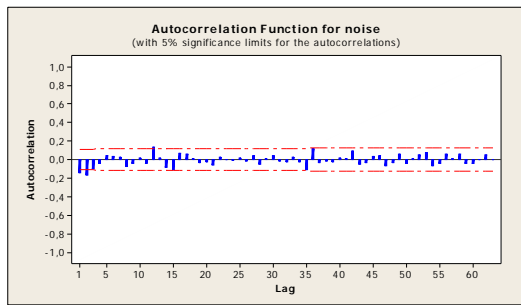


Gambar 3 Plot korelasi silang

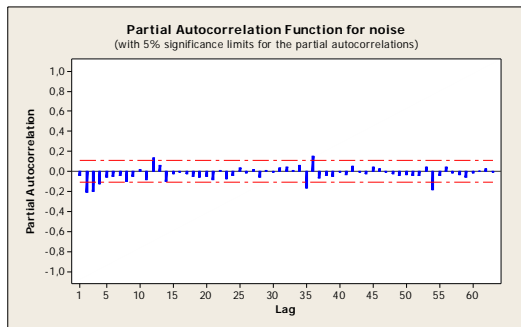
Berdasarkan hasil korelasi silang nilai yang signifikan adalah pada *lag* 1 setelah *lag* 0, sehingga parameter b bernilai 1, artinya terjadi satu hari penundaan sebelum curah hujan mulai mempengaruhi debit air di PLTA Tes. Sehingga diperoleh model fungsi transfer (b,r,s) adalah (1,0,2) dan (1,1,1).

#### D. Penaksiran Parameter Model Fungsi Tranfer

Setelah mengidentifikasi orde b,r,s selanjutnya menghitung nilai deret gangguan. Berikut plot ACF dan PACF deret gangguan :



Gambar 4.a Plot ACF deret gangguan



Gambar 4.b Plot PACF deret gangguan

Berdasarkan gambar 4.a dan 4.b, plot ACF terpotong pada lag 1, sehingga diperoleh model ARMA (1,1). Jadi, bentuk persamaan deret gangguan adalah sebagai berikut:

$$n_t = \frac{1 - 0,9809B}{1 - 0,7455B} a_t$$

Tabel 4 Pendugaan parameter fungsi transfer

Model	Parameter	Taksiran	P_value
(1,0,2)	$\phi_1$	0,74432	0,0001
	$\theta_1$	0,95453	0,0001
	$\omega_0$	0,1988	0,0001
	$\omega_1$	0,0049819	0,8591
(1,1,1)	$\phi_1$	0,73155	0,0001
	$\theta_1$	0,95821	0,0001
	$\omega_0$	0,23998	0,0159
	$\omega_1$	0,10442	0,0001
	$\delta_1$	0,80777	0,0001

Berdasarkan Tabel 4, terlihat bahwa parameter yang signifikan terdapat pada model (1,1,1) maka model fungsi transfer (1,1,1) yang dipilih sebagai model terbaik untuk peramalan debit air periode hari kedepan. Sehingga persamaan model fungsi transfer untuk meramalkan debit air pada PLTA Tes dengan mempertimbangkan curah hujan sebagai faktor input adalah sebagai berikut:

$$y_t = \frac{(0,23998 - 0,10442B)}{(1 - 0,80777B)} x_{t-1} + \frac{1 - 0,95821B}{1 - 0,73155B} a_t$$

Atau dapat diubah menjadi:

$$y_t = (1,53932)y_{t-1} + (0,59092)y_{t-2} + (0,23998)x_{t-1} - (0,27998)x_{t-2} + (0,07639)x_{t-3} + a_t - (0,76598)a_{t-1} + (0,77401)a_{t-2}$$

#### E. Uji Diagnosa Model

Uji diagnosa model fungsi transfer dilakukan terhadap uji korelasi silang (CCF) dan autokorelasi nilai sisa, dengan tujuan untuk menguji apakah asumsi bahwa deret input yang telah diputihkan adalah bebas dari komponen *noise*.

Tabel 5 Diagnosa model fungsi transfer (1,1,1)(1,1)

Uji	Lag	Stat Q	P_value
CCF	5	2,83	0,4183
	11	3,28	0,9520
	17	8,13	0,9184
	23	13,82	0,8771
	29	16,64	0,9395
	35	21,03	0,9471
	41	22,71	0,9827
	47	26,76	0,9859
	53	32,81	0,9776
	59	36,18	0,9858
ACF	6	6,23	0,1824
	12	15,57	0,1127
	18	20,71	0,1898
	24	22,32	0,4410
	30	24,60	0,6497
	36	30,67	0,6316
	42	40,34	0,4550
	48	47,93	0,3945
	54	56,54	0,3093
	60	60,80	0,3755

Berdasarkan uji Ljung-Box pada Tabel 5, nilai p\_value lebih besar dari 0,05. Artinya bahwa korelasi silang dari model tidak berbeda nyata dari nol, sehingga dapat disimpulkan residual memenuhi asumsi *white noise*. Dengan demikian, model fungsi transfer (1,1,1)(1,1) merupakan model yang layak digunakan sebagai peramalan.

#### F. Peramalan

Setelah dilakukan pemodelan terhadap ARIMA dan fungsi transfer dilanjutkan dengan tahap peramalan. Hasil ramalan terhadap debit air untuk 3 periode kedepan dapat dilihat pada Tabel 6.

Tabel 6 Hasil ramalan debit air berdasarkan data curah hujan

No	Periode	Tanggal	Ramalan ( $\frac{m^3}{s}$ )
1	366	01 Jan 2014	39,07
2	367	02 Jan 2014	38,89

3	368	03 Jan 2014	38,41
---	-----	-------------	-------

## V. KESIMPULAN

Hasil penelitian menunjukkan bahwa model fungsi transfer yang diperoleh dapat menjelaskan hubungan debit air PLTA dengan intensitas curah hujan satu hari sebelumnya, dengan model fungsi transfer (1,1,1)(1,1) sebagai berikut:

$$y_t = (1,53932)y_{t-1} + (0,59092)y_{t-2} + (0,23998)x_{t-1} - (0,27998)x_{t-2} + (0,07639)x_{t-3} + a_t - (0,76598)a_{t-1} + (0,77401)a_{t-2}$$

Sehingga peramalan debit air 3 periode kedepan terhadap curah hujan menghasilkan nilai ramalan pada tanggal 01 Januari 2014, 02 Januari 2014 dan 03 Januari 2014 berturut-turut  $39.07 \frac{m^3}{s}$ ,  $38.89 \frac{m^3}{s}$  dan  $38.41 \frac{m^3}{s}$ .

Untuk penelitian selanjutnya, sebaiknya mengkaji lebih jauh hubungan antara curah hujan dan debit air, misalnya dengan menggunakan metode fungsi transfer Multivariat bahkan analisis intervensi. Ada baiknya dilakukan analisis peramalan yang lebih kompleks dengan melibatkan lebih banyak variabel input karena debit air tidak hanya dipengaruhi oleh jumlah curah hujan tetapi juga dipengaruhi oleh faktor-faktor lain yang sangat berpengaruh agar mendapatkan nilai ramalan yang lebih baik.

## DAFTAR PUSTAKA

- Baros, S.A. 2009. *Pembangkitan Tenaga Air dan Aliran Sungai*. <http://tiagotetung.blogspot.com/2009/12/pembangkitan-tenaga-air-dan-aliran/>. (21 April 2014)
- Bowerman, B.L and J.W. Tukey. 1979. *Time Series and Forecasting: An Applied Approach*. Publishing Company: Boston.
- Bowerman, B.L and R.T. O'Connell. 1993. *Forecasting and Time Series*. Wadsworth: California.
- Box, G.E.P and G.M. Jenkins. 1994. *Time Series Analysis: Forecasting and Control*. Third Edition. Prentice Hall: New Jersey.
- Brockwell, P.J and R.A. Davis. 2002. *Introduction to Time Series and Forecasting*. Springer: USA.
- Cryer, J.D. 1986. *Time Series Analysis*. PWS-KENT Publishing Company: Boston.
- Firdaus, M. 2006. *Analisis Deret Waktu Satu Ragam*. IPB Press: Bogor.
- Hanke, J.E., A.G. Reitsch and D.W. Wichern. 2003. *Business Forecasting*. Sixth Edition. Prentice Hall: New Jersey.
- Haryanto, Hery. 2011. *Danau Tes Alami Pendangkalan*. <http://h2aryanto.wordpress.com/2011/01/27/danau-tes-alami-pendangkalan/>. (30 April 2014).
- Indrawati, F dan B. Sutijo. 2012. "Pemodelan Jumlah Ketersediaan Beras untuk Jawa Timur Dengan Pendekatan Fungsi Transfer". *Jurnal Sains dan Seni ITS* Vol. 1 (1)

- Makridakis, S., S.C Wheelwright dan V. McGee. Alih bahasa Ir. Hari Suminto. 1999. *Metode dan Aplikasi Peramalan*. Edisi revisi. Binarupa Aksara: Jakarta.
- Mulyana. 2004. *Analisis Data Deret Waktu*. Universitas Padjadjaran: Bandung.
- Mursanto, W.B 2010. "Penentuan Debit Air Outlet PLTA Maninjau". *Jurnal RACE* Vol. 4 (2).
- Wahid, Abdul. 2009. "Analisis Faktor-faktor yang Mempengaruhi Debit Sungai Mamasa". *Jurnal SMARTek* Vol 7.
- Wei, W.W.S. 2006. *Time Series Analysis: Univariate and Multivariate Methods*. Addison-Wesley Publishing Company: New York.

**APLIKASI METODE LINKAGE DENGAN JARAK EUCLID, KUADRAT  
EUCLID, CITY-BLOCK ATAU MANHATTAN DAN CHEBYCHEV  
MENGUNAKAN PROGRAM R 2.14.**

(Studi Kasus: Hasil Produksi Pangan Di Indonesia Tahun 2013)

**Masni Uli Br Purba<sup>1</sup>, Sigit Nugroho<sup>2</sup>, Jose Rizal<sup>2</sup>**

**1** Alumni Jurusan Matematika Fakultas MIPA Universitas Bengkulu

**2** Staf Pengajar Jurusan Matematika Fakultas MIPA Universitas Bengkulu

Email : [masniuli\\_b@yahoo.com](mailto:masniuli_b@yahoo.com)

**ABSTRACT**

Penelitian ini bertujuan untuk mengelompokkan 34 Propinsi di Indonesia berdasarkan Data Hasil Produksi Pangan di Indonesia pada Tahun 2013 dan mengkaji secara statistik peran jarak dan metode pengklasteran pada pengklasteran Hirarki. Data Hasil Produksi Pangan di Indonesia pada Tahun 2013 diperoleh dari Badan Pusat Statistik (BPS). Dalam penelitian ini digunakan analisis kluster dengan metode hierarki (*Single linkage, Complete Linkage, Average Linkage*) dengan jarak *Euclid*, *Kuadrat Euclid*, *City-Block* atau *Manhattan* dan *Chebychev*. Hasil penelitian menunjukkan bahwa jika ditentukan tiga kluster yang terbentuk dengan metode pengklasteran yang berbeda, jarak *City-Block* atau *Manhattan* akan menghasilkan pengklasteran yang sama. Dalam kasus Data Hasil Produksi Pangan di Indonesia Tahun 2013, jarak *City-Block* atau *Manhattan* merupakan jarak yang paling baik.

Kata kunci: Analisis Kluster, Metode Hierarki, *Linkage*, Jarak (*Distance*)

## **1. PENDAHULUAN**

Analisis Kluster adalah analisis statistika yang bertujuan untuk mengklasterkan data sedemikian sehingga data yang berada dalam kluster yang sama mempunyai sifat yang relatif homogen daripada data yang berada dalam kluster yang berbeda (Purnamasari, 2011). Terdapat beberapa metode pengklasteran dalam Analisis Kluster, yaitu metode Hierarki dan metode Non-Hierarki (pengklasteran *K-Means*). Pada metode Hierarki terdapat dua tipe dasar kluster yaitu *Agglomerative* (penggabungan) dan *Divisive* (pemecahan). Pengklasteran *Agglomerative* (penggabungan) terdiri dari tiga metode yaitu *Linkage*, *Ward's* dan *Centroid* (Nugroho, 2008). Pada analisis kluster ukuran kemiripan atau ketidakmiripan yang digunakan adalah jarak (*distance*). Menurut Nugroho

(2008), terdapat beberapa metode untuk mengukur jarak antara dua objek dengan ukuran kedekatan yang berbeda yaitu *Euclid*, *Kuadrat Euclid*, *City-Block* atau *Manhattan*, dan *Chebychev*.

Dalam penelitian ini akan dipelajari tentang pengklasteran hierarki khususnya metode *Linkage* (*Single Linkage, Complete Linkage* dan *Average Linkage*) menggunakan empat jarak yang berbeda yaitu *Euclid*, *Kuadrat Euclid*, *City-Block* atau *Manhattan* dan *Chebychev* dengan menggunakan Data Hasil Produksi Pangan Tahun 2013 di Propinsi yang ada di Indonesia.

## **2. LANDASAN TEORI**

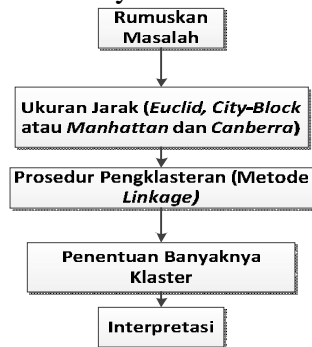
### **2.1 Analisis Kluster**

Analisis kluster merupakan salah satu analisis multivariat yang termasuk dalam metode interdependensi yaitu variabel bebas  $x$  atau faktor penyebab

tidak dibedakan dengan variabel terikat  $y$  atau respon (Nugroho, 2008). Tujuan analisis kluster adalah mengelompokkan objek-objek berdasarkan kesamaan karakteristik diantara objek-objek tersebut. Objek dapat berupa produk, benda, serta orang. Objek tersebut akan diklasifikasikan ke dalam satu atau lebih kluster sehingga objek-objek yang berada dalam satu kluster akan mempunyai kemiripan satu dengan yang lain (Santoso, 2004). Dalam analisis kluster terdapat beberapa asumsi yang harus dipenuhi yaitu sampel yang diambil benar-benar dapat mewakili populasi yang ada dan tidak multikolinieritas.

## 2.2 Proses Analisis Kluster

Menurut Supranto (2010) untuk menentukan banyaknya kluster yang akan terbentuk, perlu dilakukan proses analisis kluster. Beberapa tahap yang harus dilakukan yakni:



Gambar 1. Proses Analisis Kluster

## 2.3 Pengukuran Konsep Kemiripan Antar Objek

Menurut Nugroho (2008) pengukuran jarak untuk Mengukur jarak (*distance*) antara dua objek terdiri dari jarak *Euclid*, *Kuadrat Euclid*, *City-Block* atau *Manhattan*, dan *Chebychev*. Pada penelitian ini jarak yang akan digunakan adalah *Euclid*, *Kuadrat Euclid*, *City-Block* atau *Manhattan* dan *Chebychev* dengan persamaan yang dimiliki masing-masing jarak sebagai berikut:

### 1. Jarak *Euclid*

$$d(x, y) = \sqrt{\sum_{i=1}^p (x_i - y_i)^2}$$

### 2. Jarak kuadrat *Euclid*

$$d(x, y) = \sum_{i=1}^p (x_i - y_i)^2$$

### 3. Jarak *City-Block* atau *Manhattan*

$$d(x, y) = \sum_{i=1}^p |x_i - y_i|$$

### 4. Jarak *Chebychev*

$$d(x, y) = \max |x_i - y_i|$$

## 2.4 Metode Hierarki

Metode Hierarki adalah metode yang memulai pengelompokkan dengan dua atau lebih objek yang mempunyai kesamaan paling dekat. Menurut Nugroho (2008), metode hierarki lebih populer dibandingkan metode pengklasteran K-means (Non- Hierarki) namun metode ini membutuhkan waktu lama jika datanya besar. Dalam Metode hierarki terdapat dua pengklasteran yaitu *Agglomerative* (penggabungan) dan *Devisive* (pemecahan). Berdasarkan ukuran kemiripannya metode *Agglomerative* terdiri dari beberapa metode yaitu metode *Linkage*, metode *Ward*, metode *Centroid* dan metode *Median* (Nugroho, 2008). Metode yang akan dibahas dalam penelitian adalah metode *Linkage* (*Single linkage*, *Complete Linkage*, dan *Average Linkage*).

### 2.4.1 *Single Linkage*

Menurut Supranto (2010), metode *single linkage* adalah proses pengklasteran yang didasarkan pada jarak terdekat antar objek pada matriks *proximity* dan proses penggabungan objek-objek dilakukan berdasarkan objek-objek yang memiliki jarak terkecil (minimum) tersebut, misal  $U$  dan  $V$ , sehingga diperoleh kluster

(UV). Untuk mencari kluster antara (UV) dan kluster W lainnya dapat diperoleh dengan cara sebagai berikut:

$$d_{(UV)W} = \min\{d_{UW}, d_{VW}\}$$

$d_{UW}$  adalah anggota jarak dari kluster U ke W dan  $d_{VW}$  adalah anggota jarak dari kluster V ke W.

#### 2.4.2 Complete Linkage

Menurut Rachmatin (2014), pada metode *complete linkage* proses pengklasteran dimulai dengan menentukan jarak terdekat pada matriks *proximity* dan proses penggabungan objek-objek dilakukan berdasarkan objek-objek yang memiliki jarak terjauh (maksimum) tersebut, misal U dan V, sehingga diperoleh kluster (UV). Selanjutnya jarak antara (UV) dan kluster W lainnya dapat diperoleh dengan cara sebagai berikut:

$$d_{(UV)W} = \max\{d_{UW}, d_{VW}\}$$

$d_{UW}$  adalah anggota jarak dari kluster U ke W dan  $d_{VW}$  adalah anggota jarak dari kluster V ke W.

#### 2.4.3 Average Linkage

Menurut Rachmatin (2014), metode *average linkage* adalah metode pengklasteran yang didasarkan pada jarak rata-rata antar objek. Menurut Supranto (2010), metode *average linkage* merupakan gabungan dari *single linkage* dan *complete linkage*. Pada metode *average linkage* proses pengklasteran dimulai dengan menentukan jarak terdekat dalam matriks *proximity* dan proses penggabungan objek-objek dilakukan berdasarkan nilai rata-rata jarak dari objek-objek tersebut, misal U dan V, sehingga diperoleh kluster (UV). Selanjutnya jarak antara (UV) dan kluster W lainnya dapat diperoleh dengan cara sebagai berikut:

$$d_{(UV),W} = \frac{[\sum_{i=1}^n \sum_{k=1}^k d_{ik}]}{N_{(UV)}N_{(W)}}$$

dimana  $d_{ik}$  adalah jarak antara objek  $i$  dalam kelompok (UV) dan objek  $k$

dalam kelompok W, dan  $N_{(UV)}$  dan  $N_w$  berturut-turut adalah banyaknya objek-objek dalam kelompok (UV) dan W.

### 2.5 Interpretasi

Menurut Supranto (2010), interpretasi meliputi pengkajian mengenai centroids yaitu rata-rata nilai objek yang terdapat dalam kluster pada setiap variabel. Nilai centroid memungkinkan kita untuk menguraikan setiap kluster dengan cara memberikan suatu nama atau label.

### 3. Metode penelitian

Penelitian ini merupakan penelitian statistik terapan, yaitu suatu penelitian yang dilakukan dengan mengaplikasikan metode statistika ke dalam bidang-bidang tertentu. Pada penelitian ini diterapkan metode Analisis Kluster yaitu metode *Linkage* menggunakan tiga jarak berbeda yaitu jarak *Euclid*, Kuadrat *Euclid*, *City-Block* atau *Manhattan* dan *Chebyshev*. Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) yang telah dipublikasikan yaitu data hasil produksi pangan periode tahun 2013.

Variabel yang digunakan dalam penelitian ini adalah  $X_1$  (Hasil produksi pangan jenis tanaman Padi dalam satuan ton),  $X_2$  (Hasil produksi pangan jenis tanaman Jagung dalam satuan ton),  $X_3$  (Hasil produksi pangan jenis tanaman Kedelai dalam satuan ton),  $X_4$  (Hasil produksi pangan jenis tanaman Kacang Tanah dalam satuan ton),  $X_5$  (Hasil produksi pangan jenis tanaman Kacang Hijau dalam satuan ton),  $X_6$  (Hasil produksi pangan jenis tanaman Ubi Jalar dalam satuan ton), dan  $X_7$  (Hasil produksi pangan jenis tanaman Ubi Kayu dalam satuan ton).

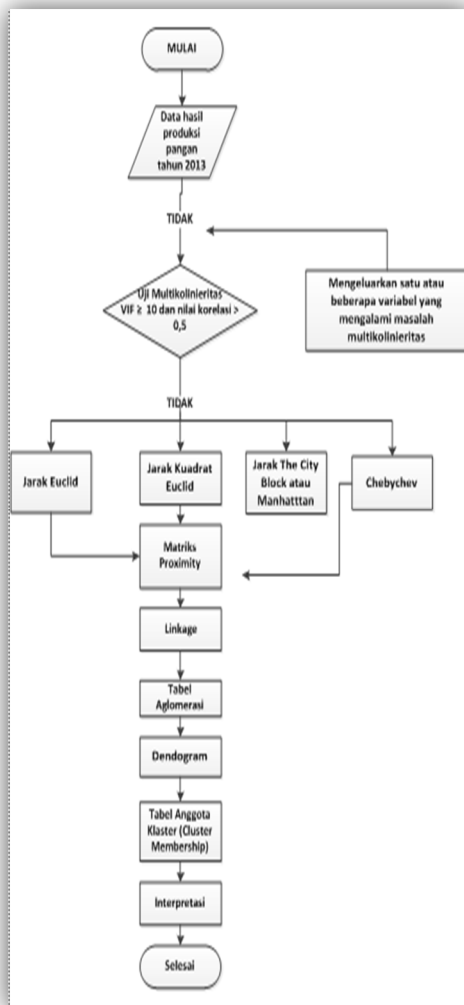
Tahapan dalam analisis dari penelitian ini adalah pertama melakukan uji asumsi yang harus dipenuhi dalam analisis kluster yaitu sampel yang diambil benar-benar dapat

mewakili populasi yang ada dan tidak multikolinieritas. Kedua mengeluarkan satu atau beberapa variabel yang mengalami masalah multikolinieritas untuk mengatasi masalah multikolinieritas pada data. Ketiga memilih ukuran jarak antar objek dengan menggunakan jarak *Euclid*, Kuadrat *Euclid*, *City-Block* atau *Manhattan* dan *Chebychev*. Keempat melakukan proses pengklasteran akan dilakukan pengelompokan data dengan mengolah data menggunakan metode *Linkage*, dan kelima melakukan interpretasi terhadap klaster yang telah terbentuk dengan cara memberi nama untuk menggambarkan karakteristik masing-masing klaster.

#### 4. HASIL PENELITIAN

##### 4.1 Asumsi Analisis Klaster

Menurut Santoso (2010), ada beberapa asumsi yang harus dipenuhi yaitu sampel yang diambil benar-benar dapat mewakili populasi yang ada dan tidak multikolinieritas. Penelitian ini menggunakan populasi sebagai total sampling. Dengan demikian asumsi pertama telah terpenuhi. Asumsi kedua yaitu tidak mengalami masalah Multikolinieritas. Pada penelitian ini untuk mengetahui ada tidaknya multikolinieritas, peneliti menggunakan uji korelasi dan uji *Variance Inflation Factor* (VIF) untuk mendeteksi masalah multikolinieritas. Berdasarkan uji korelasi dan uji *Variance Inflation Factor* (VIF) dari output program R menunjukkan bahwa variabel jagung dan kacang tanah jagung mengalami masalah multikolinieritas, untuk mengatasi masalah tersebut maka dapat diatasi dengan cara mengeluarkan satu atau beberapa variabel yang mengalami masalah multikolinieritas (Gujarati, 2004) maka dari itu variabel jagung dan kacang tanah harus dikeluarkan dari tahap analisis maka dari itu variabel padi, kedelai, kacang hijau, ubi kayu, dan ubi jalar dapat digunakan untuk tahap analisis selanjutnya.



Gambar 2. Flowchart (Diagram Alir)

##### 4.2 Proses dan Hasil Pengklasteran

Proses pengklasteran hirarki dilakukan dengan menggunakan *software* R versi 2.14.1. Metode pengklasteran hirarki dilakukan dengan memperhatikan ukuran kemiripan antar objek yang terdapat dalam matriks *proximity*. Metode pautan (*Linkage Methods*) terdiri dari pautan tunggal (*Single Linkage*), pautan lengkap (*Complete Linkage*), dan pautan rata-rata (*Average Linkage*). Pada metode ini tidak ada sumber yang mengharuskan menggunakan jarak tertentu untuk mendapatkan matriks *proximity*nya.

Untuk itulah digunakan metode jarak *Euclid*, Kuadrat *Euclid*, *City-Block* atau *Manhattan* dan *Chebychev*. Dari analisis yang dilakukan diperoleh hasil pengklasteran dengan metode pengukuran jarak *City-Block* atau *Manhattan* memberikan hasil pengklasteran yang sama persis untuk metode *Linkage*. Sedangkan pengukuran jarak *Euclid*, Kuadrat *Euclid*, dan *Chebychev* menghasilkan pengklasteran yang berbeda, namun perbedaannya tidak terlalu mencolok.

Output metode pengklasteran metode *Linkage* disajikan secara visual berbentuk dendrogram yaitu suatu bagan yang menyajikan banyaknya kluster terbesar hingga terkecil. Dendrogram dibaca dari kiri ke kanan. Garis tegak lurus menunjukkan objek yang digabung membentuk satu kluster. Sedangkan posisi garis pada skala menunjukkan jarak kluster yang digabung. Tabel Agglomerasi di bawah ini menunjukkan cara penggabungan kluster setiap tahap pada metode pautan tunggal (*Single Linkage*) dengan menggunakan jarak *Euclid*.

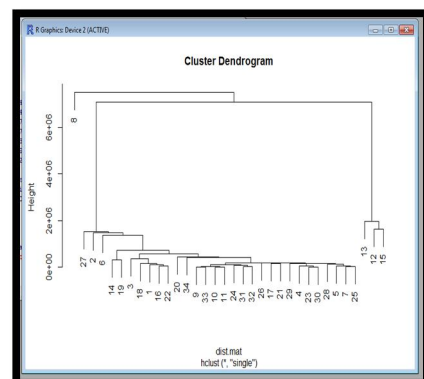
Tabel 2. Proses Agglomerative metode *Single Linkage* untuk jarak *Euclid*

Step	Clusters	Joined	Distance Level
1	23	30	1,19E+16
2	9	33	3,65E+15
3	10	11	2,51E+16
.	.	.	.
.	.	.	.
.	.	.	.
32	29	31	5,60E+19
33	32	8	5,60E+19

Pada Tahap Pertama proses pengabungan dilakukan terhadap Provinsi Kalimantan Timur (23) dan Provinsi Sulawesi Barat (30) dengan nilai jarak yang dihasilkan berdasarkan jarak *Euclid* adalah 109303,456 sehingga terbentuk kluster pada tahap pertama. Pada tahap kedua proses pengabungan dilakukan terhadap

Provinsi KEP. Bangka Belitung (9) dan Provinsi Papua Barat (33) dengan nilai jarak yang dihasilkan berdasarkan jarak *Euclid* adalah 60450,84 maka terbentuk kluster kedua dan begitu seterusnya sampai tahap ke 33.

Pada dendrogram metode pautan tunggal (*Single Linkage*) dengan jarak *Euclid*, kluster yang dapat dibentuk adalah 2 kluster, 3 kluster, 4 kluster dan seterusnya, dimana jika diambil 2 kluster maka anggota kluster tersebut adalah objek (Lampung) dan 33 objek lainnya. Jika diambil 3 kluster maka anggotanya adalah kluster dengan objek (Lampung), objek (Jawa Timur, Jawa Barat dan Jawa Tengah), dan 30 objek lainnya. Jika diambil 4 kluster maka anggotanya adalah kluster dengan objek (Lampung), objek (Jawa Tengah, Jawa Timur dan Jawa Barat), objek (Sulawesi Selatan, Sumatera Utara, Sumatera Selatan) dan 27 objek lainnya.



Gambar 3. Dendrogram Metode *Single Linkage* menggunakan Jarak *Euclid*

Jika yang diambil 3 kluster sebagai pemecahannya, maka:

- Untuk semua konsep pengukuran jarak pada metode *Complete Linkage* dan *Complete Linkage* didapat anggota pengklasteran sebagai berikut: (Jawa Timur, Jawa Barat dan Jawa Tengah), (Lampung) dan 30 objek lainnya.
- Metode *Average Linkage*, *Complete Linkage*, dan *Single Linkage*, berdasarkan konsep

pengukuran Jarak *City-Block* atau *Manhattan* anggota klasternya adalah (Jawa Timur, Jawa Barat dan Jawa Tengah), (Lampung) dan 30 objek lainnya.

- Metode *Single Linkage* dengan konsep pengukuran jarak *Euclid*, Kuadrat *Euclid* dan *Chebychev* anggota klasternya adalah objek (Lampung), objek (Jawa Timur, Jawa Barat dan Jawa Tengah), dan 30 objek lainnya.

#### 4.3 Interpretasi

Interpretasi meliputi pengkajian mengenai *centroid* yaitu rata-rata nilai objek yang terdapat dalam kluster pada tiap variabel. Nilai *centroid* memungkinkan untuk menguraikan setiap kluster dengan cara memberi suatu label atau nama. Label suatu kluster juga dapat didasarkan pada manfaat yang akan dicari (Nugroho, 2008).

Untuk metode *Single Linkage* menggunakan Jarak *Euclid Kuadrat Euclid* dan *Chebychev*, kluster pertama dapat digolongkan menjadi kluster yang memiliki Hasil Produksi Pangan yang Cukup Baik, kluster kedua dapat digolongkan menjadi kluster yang memiliki Hasil Produksi Pangan yang Kurang Baik dan kluster ketiga dapat digolongkan menjadi kluster yang memiliki Hasil Produksi Pangan yang Baik.

Untuk metode *Linkage* menggunakan Jarak *City-Block* atau *Manhattan*, kluster pertama dapat digolongkan menjadi kluster yang memiliki Hasil Produksi Pangan yang Baik, kluster kedua dapat digolongkan menjadi kluster yang memiliki Hasil Produksi Pangan yang Cukup Baik, kluster ketiga dapat digolongkan menjadi kluster yang memiliki Hasil Produksi Pangan yang Kurang Baik.

Untuk Metode *Complete Linkage* dan *Average Linkage Euclid Kuadrat Euclid City-Block* atau *Manhattan* dan *Chebychev*, kluster pertama dapat digolongkan menjadi kluster yang memiliki Hasil Produksi Pangan yang Baik, kluster kedua dapat digolongkan menjadi kluster yang memiliki Hasil Produksi Pangan yang Cukup Baik, kluster ketiga dapat digolongkan menjadi kluster yang memiliki Hasil Produksi Pangan yang Kurang Baik.

#### 5. Kesimpulan dan Saran

Pengklastran pada data Hasil Produksi Pangan di Indonesia tahun 2013 menghasilkan matriks *proximity* dan proses *agglomerative* (penggabungan) yang berbeda untuk setiap konsep pengukuran dan metode pengklastran yang digunakan. Metode *Single Linkage* memiliki anggota pengklastran dan interpretasi yang berbeda untuk ketiga jarak yang digunakan. Sedangkan pada metode *Complete Linkage* dan *Average Linkage* untuk jarak *Euclid*, Kuadrat *Euclid*, *City-Block* atau *Manhattan* dan *Chebychev* memiliki anggota pengklastran dan interpretasi yang sama. Dalam kasus Data Hasil Produksi Pangan di Indonesia Tahun 2013, jarak *City-Block* atau *Manhattan* merupakan jarak yang paling baik karena memiliki anggota pengklastran, interpretasi dan dendogram yang sama pada setiap metode pengklastran *Linkage*, disarankan untuk melanjutkan penelitian ini pada metode pengklastran nonhirarki.

## DAFTAR PUSTAKA

- Anonim. 2007. *Proyeksi Penduduk 2000-2025*. [www.bps.go.id](http://www.bps.go.id). Indonesia.
- Admaja, S. L. 2009. *Statistika Untuk Bisnis dan Ekonomi*. Andy Yogyakarta. Yogyakarta.
- Afriana, M. 2014. *Penentuan Awal Keanggotaan Klasteran Hirarki K-Means*. Skripsi, FMIPA Universitas Bengkulu. Bengkulu.
- Apreity, H. 2015. *Pengklasteran Dengan Metode Average Linkage*. Skripsi, FMIPA Universitas Bengkulu. Bengkulu.
- Carsono, N. 2009. Peran Pemuliaan dalam Meningkatkan Produksi Pertanian di Indonesia. *Faperta UNPAD*. Jatinagor.
- Dewi, S. Y. 2010. OLS, LASSO dan PLS Pada Data Mengandung Multikolinieritas. *Jurnal Ilmu Dasar* Vol. 11 No. 1, Januari 2010 :83-91.
- Ferawati, L. 2006. *Pengklasteran Provinsi Diindonesia Menurut Status Daerah Rawan Pangan Berdasarkan Indikator Yang Mempengaruhi Tahun 2006*. Skripsi, FMIPA Universitas Bengkulu. Bengkulu.
- Gujarati, D. 2010. *Ekonometri Dasar*. Selemba Empat. Jakarta.
- Hanani, N. 2012. Strategi Pencapaian Ketahanan Pangan Keluarga. *E-Journal Ekonomi Pertanian* Vol. 1, No.1.
- Izenman, A.J. 2008. *Modern Multivariate Statistical Techniques*. Springer Texts in Statistics: Philadelphia.
- Johnson, R.A. dan Wichern, D.W. 2002. *Applied Multivariate Statistical Analysis*. Prentice Hall International, New Jersey.
- Lestari, P. T. 2012. *Analisis Kluster Sekolah Menengah Atas (Sma) Berdasarkan Jumlah Tenaga Guru, Guru Sertifikasi, Pengawai Dan Siswa*. Skripsi, FMIPA Universitas Bengkulu. Bengkulu.
- Lubis, A. I. 2008. *Kajian Metode Pengklasteran Hirarki Dengan Berbagai Jarak*. Skripsi, FMIPA Universitas Bengkulu. Bengkulu.
- Mangkuatmodjo, S. 2004. *Statistika Lanjutan*. Rineka Cipta. Jakarta
- Marcus, G. L, Wattimanela, H. J. dan Lesnussa, Y. A. 2012. Analisis Regresi Komponen Utama Untuk Mengatasi Multikolinieritas Dalam Analisis Regresi Linier Berganda. *Jurnal Berekeng* Vol. 6 No. 1 Hal. 31-40 (2012).
- Mattjik, A. A dan I. M. Sumertajaya. 2011. *Sidik Peubah Ganda Dengan menggunakan SAS*. IPB: Bandung.
- Mulyana, A. 2012. Pengutan Ketahanan Pangan Untuk Menekan Jumlah Penduduk Miskin dan Rentan Pangan di Tingkat Nasional Dan Regional. *E-Journal Ekonomi Pertanian* Vol. 1, No.1.
- Nugroho, S. 2008. *Statistika Multivariat Terapan*. UNIB Press. Bengkulu.
- Oktarina. 2014. *Metode Linkage Pada Persentase Kelahiran Balita Menurut Penolong Kelahiran Terakhir*. Skripsi, FMIPA Universitas Bengkulu. Bengkulu.

- Prabowo, R. 2010. *Kebijakan Pemerintah dalam Mewujudkan Ketahanan Pangan di Indonesia. Jurnal Pertanian* **Vol. 6**, No.2.
- Purnamasari, S. M. 2011. Analisis Kelompok. *Makalah II2092 Probabilitas dan Statistik*. ITB.
- Rachmatin, D. 2014. Aplikasi Metode-Metode Agglomerative Dalam Analisis Kluster Pada Data Tingkat Polusi Udara. *Jurnal Ilmiah Program Studi Matematika STKIP Siliwangi Bandung*, **Vol 3**, No. 2, September 2014.
- Rencher, A.C., 2002. *Methods Of Multivariate Analysis Second Edition*. Wiley Series In Probability and Mathematical Statistics: Canada.
- Rosiatun, A. Widiariyah, T & Safitri, D. 2010. Analisis Kluster Untuk Segmentasi Pemirsas Program Berita Sore Stasiun TV Swasta. *Media Statistika*, **vol. 3**, No. 2, Desember 2010: 93-102.
- Santoso, S. 2010. *Statistika Multivariat Konsep & Aplikasi dengan SPSS*. Jakarta: PT. Elex Media Komputindo.
- Saputra, D. E. 2014. *Pengklasteran Hirarki Untuk Daerah Produksi Ikan Di Provinsi Bengkulu Pada Setiap Tempat Penangkapan Ikan Berdasarkan Jenis Alat Tangkapnya*. Skripsi, FMIPA Universitas Bengkulu. Bengkulu.
- Sarwono, J. 2013. *Statistik Multivariat Aplikasi untuk Riset Skripsi*. Andy Yogyakarta. Yogyakarta.
- Sari, M. N. 2012. Analisis Kluster Hirarki Dalam Pengelompokan Panti Sosial Diprovinsi Bengkulu. Skripsi, FMIPA Universitas Bengkulu. Bengkulu.
- Simatupang, J. T. 2006. Pengembangan dan Aplikasi dalam Pembangunan Pertanian di Indonesia. *Jurnal Penelitian Bidang Ilmu Pertanian* **Vol. 4**, No.1.
- Suhartono. 2008. Analisis Data Statistik Dengan R. Lab Statistik Komputasi ITS. Surabaya.
- Supranto, J. 2010. *Analisis Multivariat Arti dan Interpretasi*. Cetakan Kedua. Rineka Cipta. Jakarta.
- Suryana, N. 2011. Penggunaan Metode Statistik K-Means Clustering Pada Analisis Peruntukan Lahan Usaha Tambang Berbasis Sistem Informasi Geografi. *Statistika*, **Vol. 11**. No. 1, Mei 2011.
- Susanti, P. 2015. Pengklasteran Provinsi Diindonesia Berdasarkan Jumlah Penderita Penyakit Penyebab Kematian Tahun 2012 Dengan Menggunakan Metode K-Means. Skripsi, FMIPA Universitas Bengkulu. Bengkulu.
- Yuliana, W. N, Sukarsa, G. K. I. dan Srinadi, M. A. G. I. 2013. Perbandingan Regresi Komponen Utama Dan Robpca Dalam Mengatasi Multikolinieritas Dan Pencilan Pada Regresi Linier Berganda. *E-Jurnal Matematika* **Vol. 2**, No. 4, Nopember 2013, 1-5.
- Yulianto, S & Hidayatullah, H. K. 2014. Analisis Kluster Untuk Pengelompokan Kabupaten/Kota Diprovinsi Jawa Tengah Berdasarkan Indikator Kesejahteraan Rakyat. *Statistika*, **Vol. 2**, No.1, Mei 2014