PERBANDINGAN ANALISIS FAKTOR KLASIK DAN ROBUST BERDASARKAN BOBOT FAKTOR DAN RAGAM SPESIFIK

Skripsi

Oleh

Vinsensia Andari Kusuma Dewi

JURUSAN MATEMATIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERISTAS LAMPUNG

BANDAR LAMPUNG

2018

ABSTRAK

PERBANDINGAN ANALISIS FAKTOR KLASIK DAN ROBUST BERDASARKAN BOBOT FAKTOR DAN RAGAM SPESIFIK

Oleh

Vinsensia Andari Kusuma Dewi

Analisis faktor merupakan salah satu analisis peubah ganda yang berfungsi untuk mereduksi beberapa peubah independen menjadi lebih sedikit variabel. Terdapat dua aspek penting analisis faktor yaitu menduga bobot faktor dan ragam spesifik, dimana titik awal kerjanya adalah matriks varian kovarian atau matriks korelasi. Analisis faktor yang sering digunakan adalah analisis faktor klasik. Akan tetapi, analisis faktor klasik tidak dapat bekerja dengan baik apabila data mengandung pencilan. Agar analisis faktor tetap optimal, diperlukan adanya suatu penaksir yang *robust* yaitu analisis faktor *robust* dengan menggunakan metode *Minimum Volume Ellipsoid* (MVE) yang dapat diandalkan untuk mengatasi adanya data pencilan. Hasil penelitian ini menunjukkan bahwa analisis faktor *robust* baik digunakan untuk data pengamatan yang mengandung pencilan (pencilan 5%, 10%, 15%, 20%, 25% dan 30%), sedangkan analisis faktor klasik baik digunakan untuk data yang tidak mengandung pencilan (pencilan 0%).

Kata Kunci : Analisis Faktor, Pencilan, *Robust Minimum Volume Ellipsoid* (MVE)

ABSTRACT

COMPARATION OF CLASSIC AND ROBUST FACTOR ANALYSIS BASED ON FACTOR LOADING AND SPECIFIC VARIANCE

By

Vinsensia Andari Kusuma Dewi

Factor analysis is one of multivariate analysis that is used to reduce some independent variables into fewer variables. There are two important aspects of factor analysis, i.e. to estimate factor loading and specific variance, where the starting point of the work is the variance covariance matrix or the correlation matrix. Factor analysis that is often used is the classical factor analysis. However, the classical factor analysis can not work properly if the data contained outliers. In order for factor analysis to remain optimal, we need a robust estimator namely robust factor analysis using the Minimum Volume Ellipsoid (MVE) method that can be relied upon to overcome the existence of the outliers. The results of this study indicate that robust factor analysis is well used for observational data that contain outlier (5%, 10%, 15%, 20%, 25% and 30% outliers), whereas classical factor analysis is well used for observational data which does not contain outlier (0% outlier).

Keywords: Factor Analysis, Outlier, Robust Minimum Volume Ellipsoid (MVE)

PERBANDINGAN ANALISIS FAKTOR KLASIK DAN ROBUST BERDASARKAN BOBOT FAKTOR DAN RAGAM SPESIFIK

Oleh
Vinsensia Andari Kusuma Dewi

Skripsi

Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA SAINS

Pada
Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2018

Judul Skripsi : **PERBANDINGAN ANALISIS FAKTOR KLASIK DAN ROBUST BERDASARKAN BOBOT FAKTOR DAN RAGAM SPESIFIK**

Nama Mahasiswa : *Vinsensia Andari Kusuma Dewi*

Nomor Pokok Mahasiswa : 1417031123

Program Studi : Matematika

Fakultas : Matematika dan Ilmu Pengetahuan Alam

MENYETUJUI

1. Komisi Pembimbing

Dr. Khoirin Nisa, M.Si.	**Drs. Nusyirwan, M.Si**
NIP 19740726 200003 2 001	NIP 19661010 199205 1 001

2. Ketua Jurusan Matematika

Prof. Dra. Wamiliana, M.A, Ph.D.
NIP 19631108 198902 2 001

MENGESAHKAN

1. Tim Penguji

Ketua : **Dr. Khoirin Nisa, M.Si.** .....

Sekretaris : **Drs. Nusyirwan, M.Si** .....

Penguji
Bukan Pembimbing : **Drs. Eri Setiawan, M.Si.** .....

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Prof. Warsito, S.Si., DEA., Ph.D.
NIP 19710212 199512 1 001

Tanggal Lulus Ujian Skripsi : 2 Mei 2018

PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama	: Vinsensia Andari Kusuma Dewi
Nomor Pokok Mahasiswa	: 1417031123
Program Studi	: Matematika
Jurusan	: Matematika

Dengan ini menyatakan bahwa skripsi saya yang berjudul **"Perbandingan Analisis Faktor Klasik dan Robust Berdasarkan Bobot Faktor dan Ragam Spesifik"** adalah hasil pekerjaan saya sendiri. Semua hasil tulisan dalam skripsi ini telah mengikuti kaidah penulisan karya ilmiah Universitas Lampung. Apabila kemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau telah dibuat orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku,

Bandar Lampung, Mei 2018

Penulis

Vinsensia Andari K.D

NPM. 1417031123

RIWAYAT HIDUP

Penulis dilahirkan di Bandar Lampung tepatnya pada tanggal 9 April 1996, sebagai putri pertama dari pasangan Bapak Yohanes Sudaryadi dan Ibu Fransiska Arimbi.

Penulis telah menempuh pendidikan di Sekolah Dasar (SD) di SD Fransiskus 1 Tanjung Karang Pusat pada tahun 2002 – 2008, Sekolah Menengah Pertama (SMP) di SMP Fransiskus 1 Tanjung Karang Pusat pada tahun 2008-2011, dan Sekolah Menengah Atas (SMA) di SMA Negeri 12 Bandar Lampung pada tahun 2011-2014.

Pada tahun 2014 penulis terdaftar sebagai Mahasiswa Program Sudi S1 Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur SNMPTN. Selama menjadi mahasiswa, penulis bergabung di Generasi Muda Matematika (GEMATIKA) dan Himpunan NATURAL yang diamanahkan sebagai Anggota Redaksi periode 2015 serta Sekretaris Penelitian dan Pengembangan periode 2016.

Pada bulan Januari 2017 penulis melaksanakan Kerja Praktik (KP) di Badan Pusat Statistik Provinsi Lampung guna menerapkan ilmu yang telah diperoleh sewaktu kuliah. Pada bulan Juli 2017 penulis melaksanakan Kuliah Kerja Nyata (KKN) di Desa Bakauheni, Kecamatan Bakauheni, Kabupaten Lampung Selatan.

KATA INSPIRASI

" Do your best until you feel that's enough "
Lakukan yang terbaik sampai kamu merasa itu cukup

Kurangi berkata "nanti saja" atau "masih banyak waktu"
ketika pekerjaan atau hal lain datang menghampirimu.
Ingatlah bahwa waktu tidak berdiam diri.
Gunakan waktumu dengan bijak.

" Be thankful for what you're now and keep fighting for what you want to be tomorrow. Don't forget always pray in every your step."

PERSEMBAHAN

Puji Syukur kepada Tuhan Yang Maha Esa karena atas rahmat dan hidayah Nya, masih memberikanku kesempatan untuk menuntut ilmu di Universitas Lampung sampai saat ini.

Teruntuk ayah dan ibuku yang selalu ada disetiap perjalanan hidupku, yang memberikan doa, nasihat, semangat, dukungan baik moril maupun materil dan juga kasih sayang serta pengorbanan yang tak tergantikan. Dengan kerendahan hati yang tulus dan rahmat dari-Nya, kupersembahkan skripsi ini untuk ayah dan ibu yang selalu ada dalam doaku.

Dan untuk kedua adikku terkasih yang selalu memberikan warna dalam hidup, membuatku belajar arti kesabaran, juga memberikanku semangat ketika lelah menghampiri.

Terimakasih juga untuk saudara, sahabatku, juga teman-teman yang selalu memberikan semangat dan juga membantuku dalam proses pengerjaan skripsi ini.

“ Without my family and people around me, i’m nothing.”

“ But with them, i can be something”

SANWACANA

Puji dan syukur senantiasa kepada Tuhan Yang Maha Esa yang selalu memberikan nikmat, rahmat, serta hidayah-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul "**PERBANDINGAN ANALISIS FAKTOR KLASIK DAN ROBUST BERDASARKAN BOBOT FAKTOR DAN RAGAM SPESIFIK"**

Penulis juga mengucapkan terima kasih kepada pihak-pihak yang telah membantu dalam memberikan bimbingan, motivasi, semangat dan saran yang membangun selama proses penyusunan skripsi ini. Terimakasih kepada:

1. Ibu Dr. Khoirin Nisa, M.Si, selaku Dosen Pembimbing 1 yang telah meluangkan waktu dan membimbing penulis selama menyusun skripsi.
2. Bapak Drs. Nusyirwan, M.Si, selaku Dosen Pembimbing 2 yang telah memberikan saran serta arahan kepada penulis.
3. Bapak Drs. Eri Setiawan, M.Si, selaku Dosen Pembahas yang memberi saran dan evaluasi kepada penulis selama penyusunan skripsi.
4. Bapak Subian Saidi, S.Si., M.Si, selaku Dosen Pembimbing Akademik yang telah memberikan pengarahan selama masa perkuliahan.
5. Ibu Prof. Dra. Wamiliana, MA, Ph.D, selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam.

6. Bapak Prof. Warsito, S.Si., DEA., Ph.D, selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Dosen, staf dan karyawan Jurusan Matematika FMIPA Universitas Lampung yang telah memberikan ilmu serta bantuan kepada penulis.
8. Orang tua dan adik-adik penulis yang juga telah memberikan motivasi, doa, dukungan serta menyediakan sarana dalam penyusunan skripsi ini.
9. Sahabat serta teman-teman yang telah membantu dan memberikan dukungan kepada penulis
10. Seluruh pihak yang telah membantu penulis dalam menyusun skripsi ini.

Penulis juga menyadari bahwa dalam penulisan skripsi ini masih banyak terdapat kekurangan. Oleh karena itu, penulis mengharapkan saran dan kritik yang membangun guna penelitian selanjutnya agar menjadi lebih baik lagi.

Bandar Lampung, Mei 2018
Penulis,

Vinsensia Andari K.D

DAFTAR ISI

Halaman

DAFTAR TABEL

DAFTAR GAMBAR

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah 1

1.2 Tujuan Penelitian 3

1.3 Batasan Masalah 3

1.4 Manfaat Penelitian 3

II. TINJAUAN PUSTAKA

2.1 Konsep Matriks 4

2.1.1 Definisi Matriks 4

2.1.2 Matriks Invers 4

2.1.3 Matriks Definit dan Semidefinit Positif 5

2.1.4 Matriks Ortogonal 6

2.1.5 Matriks Simetrik 6

2.1.6 Matriks Transpose A 6

2.1.7 Matriks Identitas 7

2.1.8 Determinan Matriks 7

2.2 Analisis Multivariat 7

2.3 Analisis Faktor 9

2.3.1 Definisi Analisis Faktor 9

2.3.2 Model Faktor 10

2.4 Metode Pendugaan Bobot Faktor 12

2.4.1 Metode Komponen Utama 12
2.4.2 Metode Faktor Utama 15
2.4.3 Metode Kemungkinan Maksimum 16
2.4.4 Metode Kuadrat Terkecil Tak Terboboti 17
2.5 Pencilan 18
2.6 *Robust* 19
2.7 Metode *Minimum Volume Ellipsoid (MVE)* 20
2.8 Algoritma Resampling 22

III. METODOLOGI PENELITIAN
3.1 Waktu dan Tempat Penelitian 25
3.2 Data Penelitian 25
3.3 Metode Penelitian 26

IV. HASIL DAN PEMBAHASAN
4.1 Hasil Simulasi Analisis Faktor Klasik dan *Robust* 28
4.2 Hasil Analisis Data Bangkitan dengan Nilai *Mean* dan Matriks *Variance Covariance* Random 33
4.3 Grafik Perbandingan Nilai *Mean Square Error* Analisis Faktor Klasik dan *Robust* 35
4.3.1 Gambar Grafik Nilai MSE Tabel 1 35
4.3.2 Gambar Grafik Nilai MSE Tabel 2 37
4.4 Pembahasan 41

V. KESIMPULAN

DAFTAR PUSTAKA

LAMPIRAN

DAFTAR TABEL

Tabel Halaman

1. Perbandingan Nilai MSE Bobot Faktor serta Ragam Spesifik Klasik dan *Robust* untuk seluruh Presentase Pencilan 32
2. Perbandingan Nilai MSE Bobot Faktor serta Ragam Spesifik Klasik dan *Robust* untuk seluruh Presentase Pencilan dan 5 Gugus Data Bangkitan 33

DAFTAR GAMBAR

Gambar Halaman

1. Grafik Nilai MSE Bobot Faktor Klasik dan *Robust* pada Tabel 1 35
2. Grafik Nilai MSE Ragam Spesifik Klasik dan *Robust* pada Tabel 1 36
3. Grafik Nilai MSE Bobot Faktor Klasik dan *Robust* untuk Data Pertama 37
4. Grafik Nilai MSE Bobot Faktor Klasik dan *Robust* untuk Data Kedua 37
5. Grafik Nilai MSE Bobot Faktor Klasik dan *Robust* untuk Data Ketiga 38
6. Grafik Nilai MSE Bobot Faktor Klasik dan *Robust* untuk Data Keempat 38
7. Grafik Nilai MSE Bobot Faktor Klasik dan *Robust* untuk Data Kelima 38
8. Grafik Nilai MSE Ragam Spesifik Klasik dan *Robust* Data Pertama 39
9. Grafik Nilai MSE Ragam Spesifik Klasik dan *Robust* Data Kedua 39
10. Grafik Nilai MSE Ragam Spesifik Klasik dan *Robust* Data Ketiga 39
11. Grafik Nilai MSE Ragam Spesifik Klasik dan *Robust* Data Keempat 40
12. Grafik Nilai MSE Ragam Spesifik Klasik dan *Robust* Data Kelima 40

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Analisis faktor merupakan salah satu analisis peubah ganda yang berfungsi untuk mereduksi beberapa peubah independen menjadi lebih sedikit variabel. Dalam pendekatannya, analisis faktor yang sering digunakan adalah analisis faktor klasik dengan dua metode dasar analisis yaitu *principal component analysis* dan *common factor analysis.* *Principal component analysis* bertujuan untuk mengetahui jumlah faktor minimal yang dapat diekstrak, sedangkan *common factor analysis* bertujuan untuk mengetahui dimensi-dimensi laten atau konstruk yang mendasari variabel-variabel asli.

Dalam analisis faktor, terdapat dua aspek penting yaitu menduga bobot faktor dan ragam spesifik. Misalkan $X_1, X_2, \dots, X_n$ adalah sampel acak dari sebaran ganda p dengan E(X) = $\boldsymbol{\mu}$ dan cov(X) = $\boldsymbol{\Sigma}$. Titik awal kerja pada analisis faktor adalah matriks varian kovarian sampel atau matriks korelasi sampel dari data yang didapatkan kemudian menguraikan matriks ini untuk memperoleh dugaan bobot faktor dan ragam spesifik.

Analisis faktor klasik sangat rentan terhadap adanya pencilan. Hal ini dikarenakan rata-rata sampel dan matriks varian kovarian sampel sangat sensitif

terhadap pencilan dan berpengaruh pada keakuratan hasil analisis faktor. Agar analisis faktor tetap optimal dalam penganalisisan meskipun dalam kondisi data yang mengandung pencilan maka diperlukan adanya suatu penaksir yang *robust* (tegar terhadap pencilan). Analisis faktor yang mengandung penaksir *robust* selanjutnya disebut sebagai analisis faktor *robust*.

Salah satu metode yang *robust* terhadap pencilan adalah menduga matriks kovarian dengan menggunakan metode *Minimum Volume Ellipsoid* (MVE). MVE terkenal karena kepekaannya terhadap pencilan yang membuatnya dapat diandalkan untuk mendeteksi pencilan. Metode ini sangat berguna dalam mendeteksi pencilan pada data multivariat, karena memenuhi jumlah batas maksimum pencilan yaitu 50%.

Penduga MVE diperoleh dengan mencari subsampel berukuran h dari data berukuran n ($h < n$) yang memiliki volume ellipsoid terkecil. Nilai tengah dari MVE diduga dengan titik tengah ellipsoid dan matriks kovarian diduga dengan matriks pembentuk dari ellipsoid itu sendiri. Dalam perhitungannya, penduga MVE dihitung dengan mengunakan pendekatan algoritma resampling.

Berdasarkan penjelasan diatas, penulis tertarik untuk mengkaji mengenai analisis faktor *robust* menggunakan metode *minimum volume ellipsoid* (MVE) dalam mengatasi adanya masalah pencilan pada data multivariat.

1.2 Tujuan Penelitian

Adapun tujuan dari penulisan skripsi ini, antara lain untuk :

a. Mengkaji analisis faktor *robust* menggunakan metode *minimum volume ellipsoid* (MVE) dalam mengatasi adanya masalah pencilan pada data multivariat.

b. Memperoleh hasil evaluasi penanganan pencilan pada analisis faktor *robust* menggunakan metode MVE dan membandingkannya dengan analisis faktor klasik dalam mengatasi adanya pencilan.

1.3 Batasan Masalah

Agar tidak memperluas pembahasan maka penelitian ini dibatasi pada hal-hal berikut :

a. Data yang digunakan adalah data simulasi yang dibangkitkan menggunakan *software SAS*.

b. Metode yang digunakan dalam analisis faktor *robust* yaitu metode *minimum volume ellipsoid* (MVE).

1.4 Manfaat Penelitian

Adapun manfaat dari penelitian yang telah dilakukan antara lain :

a. Menambah pengetahuan serta wawasan dalam memahami pengembangan konsep statistika multivariat, khususnya analisis faktor.

b. Memberikan informasi kepada pembaca maupun peneliti lain mengenai metode *minimum volume ellipsoid (MVE)* yang digunakan pada analisis faktor untuk mengatasi pencilan pada data multivariat

II. TINJAUAN PUSTAKA

2.1 Konsep Matriks

2.1.1 Definisi Matriks

Sebuah matriks adalah jajaran empat persegi panjang dari bilangan-bilangan. Bilangan-bilangan dalam jajaran tersebut disebut entri dari matriks (Anton & Rorres, 2004).

Matriks dapat dituliskan sebagai berikut :

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pp} \end{bmatrix}$$

2.1.2 Matriks Invers

Invers matriks **A** merupakan matriks kebalikan dari **A**, hal tersebut dapat dinyatakan dengan simbol $\mathbf{A}^{-1}$**.** Adapun formulasi invers dinyatakan sebagai berikut :

$$\boldsymbol{A}^{-1} = \frac{1}{|\boldsymbol{A}|}\ adj\ (\boldsymbol{A})$$

dimana :

$|\boldsymbol{A}|$ = determinan **A**

$adj\ (\boldsymbol{A})$ = adjoint **A** = transpose dari matriks kofaktor.

Invers matriks **A** merupakan kebalikan dari matriks **A** nya, maka hasil dari perkalian antara matriks A dengan invernya akan menghasilkan Identitas (**I**).

$$\boldsymbol{A^{-1}A = AA^{-1} = I}$$

dimana,

$\boldsymbol{A^{-1}}$: invers matriks

A : matriks **A**

I : matriks Identitas

(Anton & Rorres, 2004).

2.1.3 Matriks Definit dan Semidefinit Positif

Sebuah matriks berukuran n x n, **A** dikatakan bersifat definit positif jika untuk sembarang vektor $\boldsymbol{x} \neq 0$, bentuk kuadratik $\boldsymbol{x^{-1}Ax} > 0$. Sedangkan dikatakan semidefinit jika $\boldsymbol{x^{-1}Ax} \geq 0$. Jika **A** adalah definit positif maka persamaan $\boldsymbol{x^{-1}Ax} = c$ dengan c adalah konstanta. Jika **A** adalah matriks diagonal yang semua unsur diagonalnya bernilai positif, maka **A** adalah matriks definit positif. Akan tetapi, apabila ada sebuah unsur bernilai nol dan yang lain positif maka **A** akan menjadi matriks semidefinit positif. Matriks diagonal yang unsurnya ragam peubah akan bersifat demikian karena ragam tidak pernah bernilai negatif.

Matriks yang definit positif, akar-akar cirinya semua bernilai positif dan determinan dari matriks definit positif juga bernilai positif karen berupa hasil perkaliannya. Jadi determinannya tidak sama dengan nol, sehingga **A** bersifat non-singular.

Matriks **B** berukuran m x n, maka $\mathbf{BB^T}$ dan $\mathbf{B^TB}$ bersifat semidefinit positif. Jika $m < n$ dan r(**B**) = m, maka $\mathbf{BB^T}$ definit positif, tetapi $\mathbf{B^TB}$ masih semidefinit positif.

2.1.4 Matriks Ortogonal

Suatu matriks bujur sangkar **A** disebut ortogonal jika $\mathbf{AA^T = A^TA = I}$ yaitu jika $\mathbf{A^T = A^{-1}}$

2.1.5 Matriks Simetrik

Sebuah matriks **A** berukuran n x n dikatakan simetrik jika $\mathbf{A^T = A}$. Misalnya, jika a_{ij} adalah unsur ke (i,j) dari matriks, maka untuk matriks simetrik adalah $a_{ij} = a_{ji}$, untuk semua i dan j.

2.1.6 Matriks Transpose A

Jika **A** adalah sebarang matriks m x n, maka transpose dari **A** dinyatakan oleh $\mathbf{A^T}$ dan didefinisikan dengan matriks n x m yang didapatkan dengan mempertukarkan baris-baris dan kolom-kolom dari **A**, sehingga kolom pertama dari $\mathbf{A^T}$ adalah baris pertama dari **A,** kolom kedua dari $\mathbf{A^T}$ adalah baris kedua dari **A**, dan seterusnya (Anton & Rorres, 2004).

2.1.7 Matriks Identitas

Jika **R** adalah bentuk eselon baris tereduksi dari matriks $\boldsymbol{A_{nxn}}$, maka terdapat dua kemungkinan yaitu **R** memiliki satu baris bilangan nol atau **R** merupakan matriks identitas $\mathbf{I_n}$ (Anton & Rorres, 2004).

2.1.8 Determinan Matriks

Menurut Mattjik dan Sumertajaya (2011), determinan dari matriks **A** berukuran n x n adalah perkalian dari semua akar ciri $A, \lambda_1, \lambda_2, \ldots, \lambda_n$ dapat dinotasikan $|\boldsymbol{A}|$**,** sehingga

$$|\boldsymbol{A}| = \lambda_n \, x \ldots x \, \lambda_n$$

Jadi, $|\boldsymbol{A}| = 0$ jika dan hanya jika paling tidak ada satu akar yang nol, yaitu terjadi jika dan hanya jika **A** singular.

2.2 Analisis Multivariat

Metode untuk menganalisis data yang terdiri lebih dari satu peubah secara simultan dikenal sebagai analisis multivariat. Apabila data yang dikumpulkan dalam suatu penelitian terdiri dari sejumlah unit objek yang besar dan pada setiap objek banyak terdapat variabel yang diukur, maka untuk menganalisis data tersebut diperlukan adanya statistik multivariat karena statistik univariat tidak lagi dapat menyelesaikan masalah secara baik.

Suatu matriks acak $\boldsymbol{X} = (X_1, X_2, \dots, X_P)$ berderajat p dikatakan berdistribusi normal multivariat dengan vektor nilai tengah **μ** dan matriks kovarian **Σ** dituliskan:

$$X \sim N_p\ (\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

Misalkan $X_1, X_2, \dots, X_P$ variabel acak dari distribusi normal multivariat dengan vektor nilai tengah **μ** dan matriks kovarian **Σ,** penduga **μ** diberikan oleh :

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = E(\boldsymbol{X}) \begin{bmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_p) \end{bmatrix} \qquad \textbf{(2.1)}$$

dengan,

$$\mu_i = \frac{1}{n}\left(\sum_{i-1}^{n} x_i\right) = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

dan matriks kovarian **Σ** didefinisikan oleh :

$$\boldsymbol{\Sigma} = Cov\ (X) = E(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})^{\ T} \qquad \textbf{(2.2)}$$

Sedangkan penduga **Σ** diberikan oleh :

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{n-1}\left[\sum_{r=1}^{n} (X_{ir} - \bar{X}_I)\left(X_{jr} - \bar{X}_J\right)\right]$$

Konsep kovarian dirangkum dalam suatu matriks yang memuat varian dan kovarian sebagai berikut :

$$\boldsymbol{\Sigma} = \begin{bmatrix} var\ (X_1) & cov\ (X_1, X_2) & \cdots & cov\ (X, X_1) \\ cov\ (X_2, X_1) & var\ (X_2) & \cdots & cov\ (X_2, X_p) \\ \vdots & \vdots & \ddots & \vdots \\ cov\ (X_p, X_1) & cov\ (X_p, X_2) & \cdots & var\ (X_p) \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix}$$

Dengan mengartikan $cov(x_i, x_i) = var(x_i) = \sigma_{ii}$ bentuk $cov(x_i, x_j)$ disebut sebagai unsur ke-(i,j) dari matriks **Σ**.

dimana:

$$\sigma_{ii} = var(x_i) = E[(x_i - \mu_i)^2] \qquad \textbf{(2.3)}$$

dan

$$\sigma_{ij} = cov(x_i, x_j) = E[(x_i - \mu_i)(x_j - \mu_j)] \qquad \textbf{(2.4)}$$

Hubungan linear antara $x_i, \dots, x_p$ dapat dicari melalui koefisien korelasi. Korelasi antara x_i dan x_j diperoleh melalui :

$$\rho_{ij} = \frac{cov(x_i, x_j)}{\sqrt{var\ (x_i)\ var(x_j)}} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\ \sigma_{jj}}} \qquad (\mathbf{2.5})$$

dalam bentuk matriks didefinisikan sebagai berikut :

$$\boldsymbol{\rho} = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & \rho_{pp} \end{bmatrix} = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & 1 \end{bmatrix}$$

2.3 Analisis Faktor

2.3.1 Definisi Analisis Faktor

Analisis faktor adalah salah satu analisis yang banyak digunakan pada statistik peubah ganda. Analisis ini pertama kali diperkenalkan oleh Spearman (1904), dan dikembangkan Thurstone (1947), Thomson (1951), Lawley (1940,1941) serta tokoh lainnya. Pada awalnya analisis ini tergolong sulit dan kontroversial, namun

dalam perkembangannya dirasakan menjadi alat yang berguna. Terutama setelah perkembangan komputer dan paket-paket piranti lunak statistik serta pada tahun 1970-an banyak terbit buku dan publikasi lain yang membahas penerapannya diberbagai bidang, seperti biologi, kimia, ekologi, ekonomi, pendidikan, ilmu politik, psikologi dan sosiologi (Mattjik, 2011).

Menurut Suliyanto (2005), analisis faktor merupakan suatu teknik untuk menganalisis tentang ketergantungan dari beberapa variabel secara simultan dengan tujuan untuk menyederhanakan dari bentuk hubungan antara beberapa variabel yang diteliti menjadi sejumlah faktor yang lebih sedikit dari variabel yang diteliti. Hal ini berarti, analisis faktor dapat juga menggambarkan tentang struktur data dari suatu penelitian.

2.3.2 Model Faktor

Misalkan **X** vektor acak dengan vektor rata-rata $\boldsymbol{\mu}_{p\ x\ 1}$ dan matriks kovarian $\boldsymbol{\Sigma}$, dan hubungan antar unsur vektor **X** dapat dituliskan dalam model faktor :

$$\underset{(p\ x\ l)}{\boldsymbol{X}-\boldsymbol{\mu}} = \underset{(p\ x\ k)}{\boldsymbol{L}}\ \underset{(k\ x\ l)}{\boldsymbol{F}} + \underset{(p\ x\ l)}{\boldsymbol{\varepsilon}} \qquad (\mathbf{2.6})$$

Persamaan (2.6) merupakan model faktor orthogonal dengan $\boldsymbol{\mu}$ adalah vektor konstanta, $\boldsymbol{F}$ adalah vektor acak dengan unsur $F_1,\ldots,F_k$, dan disebut faktor bersama, $\boldsymbol{L}$ adalah matriks konstanta yang tidak diketahui nilainya berukuran $p\ x\ k$, disebut bobot faktor, dan unsur-unsur $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_p$ adalah unsur vektor acak $\boldsymbol{\varepsilon}$ yang disebut faktor khusus.

Model faktor orthogonal dengan *m* faktor bersama dinyatakan sebagai :

$$\underset{(p\ x\ l)}{\boldsymbol{X}} = \underset{(p\ x\ l)}{\boldsymbol{\mu}} + \underset{(p\ x\ k)}{\boldsymbol{L}} \underset{(k\ x\ l)}{\boldsymbol{F}} + \underset{(p\ x\ l)}{\boldsymbol{\varepsilon}} \qquad (\mathbf{2.7})$$

Dalam hal ini diasumsikan bahwa vektor $\boldsymbol{F}$ dan $\boldsymbol{\varepsilon}$ saling tidak berkorelasi. Berimplikasi model (2.7) untuk unsur X tertentu, misalkan X_i yang mewakili pengukuran pada peubah tertentu dapat dituliskan dalam bentuk kombinasi linier dari seluruh faktor bersama dan sebuah faktor khusus ε_i ditulis sebagai berikut :

$$\begin{array}{lllllll} X_1 - \mu_1 = & l_{11}\,F_1 + & l_{12}\,F_2 + & \cdots + & l_{1m}\,F_m & \varepsilon_1 \\ X_2 - \mu_2 = & l_{21}\,F_1 + & l_{22}\,F_2 + & \cdots + & l_{2m}\,F_m & \varepsilon_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X_p - \mu_p = & l_{p1}\,F_1 + & l_{p2}\,F_2 + & \cdots + & l_{3m}\,F_m & \varepsilon_p \end{array}$$

dengan :

μ_i = rata-rata variabel asal ke-i

ε_i = faktor khusus ke-i

F_j = faktor umum ke-j

l_{ij} = bobot faktor dari variabel asal ke-i pada faktor ke-j

Struktur kovarians untuk model faktor umum adalah sebagai berikut :

1. $\text{Cov}\,(\mathbf{X}) = \mathbf{LL}^{\mathrm{T}} + \boldsymbol{\psi}$

   atau $\text{Var}(X_i) = l_{\mathrm{i1}}^2 + l_{\mathrm{im}}^2 + \psi_i$

   $\text{Cov}\,(X_i, X_k\,) = l_{\mathrm{i1}} l_{\mathrm{k1}} + l_{im}\, l_{km}$

2. $\text{Cov}\,(\mathbf{X}, \mathbf{F}) = \mathbf{L}$

   atau $\text{Cov}\left(X_i, F_j\,\right) = l_{\mathrm{ij}}$

(Johnson & Wichern, 2007)

2.4 Metode Pendugaan Bobot Faktor

2.4.1 Metode Komponen Utama

Metode komponen utama bertujuan untuk menaksir parameter pada analisis faktor, yaitu varians spesifik (Ψ_{pxp}), komunalitas (h), dan matriks bobot faktor (L_{pxk}). Matriks ragam peragam dari sampel yaitu **S** yang merupakan estimator (penduga) bagi matriks ragam peragam populasi yang tidak diketahui yaitu **Σ**. Komponen utama analisis faktor pada matriks ragam peragam **Σ** memiliki pasangan nilai eigen dan vektor eigen $(\widehat{\lambda_i}, \widehat{e_i}), i = 1,2, \dots, p$ dimana $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$ dan e_i ternormalisasi.

Dalam penentuan jumlah faktor, sebenarnya dapat diperoleh faktor sebanyak variabel yang ada, akan tetapi karena tujuan dari melakukan analisis faktor adalah mencari variabel baru yang disebut faktor yang saling tidak berkorelasi, bebas satu sama lain, lebih sedikit jumlahnya dari variabel asli, namun dapat menyerap dan menerangkan sebagian besar informasi yang terkandung dalam variabel asli. Beberapa prosedur yang bisa digunakan untuk menentukan jumlah faktor yaitu penentuan secara *apriori* (ditentukan terlebih dahulu faktornya), atau berdasarkan pada nilai eigen , *scree plot, significance test, split-half reliabily* dan *percentage of variance accounted for.* Teknik yang paling umum digunakan dalam menentukan jumlah faktor adalah dengan melihat kriteria nilai eigen . Faktor dengan nilai eigen lebih besar dari satu yang akan dipertahankan, sedangkan jika lebih kecil dari satu maka faktor tidak diikutsertakan dalam model.

Misalkan **R** adalah matriks korelasi contoh berukuran *p x p* yang simetrik dan definit positif, sehingga dapat dituliskan sebagai berikut :

$$\mathbf{R} = \mathbf{\Gamma}\,\mathbf{\Lambda}\,\mathbf{\Gamma}^{\mathbf{T}}$$

Dengan $\mathbf{\Lambda}$ adalah diag$(\lambda_1, \dots, \lambda_p)$ dan $\lambda_1 \geq \cdots \geq \lambda_p > 0$ adalah akar ciri matriks **R**, serta $\mathbf{\Gamma}\mathbf{\Gamma}^{\mathbf{T}} = \mathbf{\Gamma}^{\mathbf{T}}\mathbf{\Gamma} = \mathbf{I}_{\mathrm{P}}$**,** dengan $\mathbf{\Gamma}$ adalah matriks ortogonal *p x p* yang kolom-kolomnya adalah vektor ciri matriks **R**, yaitu $\mathbf{\Gamma}_1, \dots, \mathbf{\Gamma}_p$ yang berpadanan dengan vektor ciri $\lambda_1, \dots, \lambda_p$. Misalkan *k* adalah banyaknya komponen utama yang dipilih menggunakan kriteria tertentu, misalnya banyaknya komponen utama minimum yang mampu menerangkan presentase keragaman total. Dengan mendefinisikan matriks $\hat{\boldsymbol{L}}$ berukuran *p x k* sebagai :

$$\hat{\boldsymbol{L}} = \left[\sqrt{\lambda_1}\Gamma_1, \sqrt{\lambda_2}\Gamma_2, \dots, \sqrt{\lambda_k}\Gamma_{\mathrm{k}}\ \right] \qquad (\mathbf{2.8})$$

Maka **R** didekati dengan :

$$\hat{\boldsymbol{L}}\hat{\boldsymbol{L}}^{\boldsymbol{T}} = \sum_{i=1}^{k} \lambda_i\ \Gamma_i\ \Gamma_i{}'$$

$$\hat{\boldsymbol{L}}\hat{\boldsymbol{L}}^{\boldsymbol{T}'} = \left[\sqrt{\lambda_1}\Gamma_1, \sqrt{\lambda_2}\Gamma_2, \dots, \sqrt{\lambda_k}\Gamma_{\mathrm{k}}\ \right] \begin{bmatrix} \sqrt{\lambda_1}\Gamma_1 \\ \sqrt{\lambda_2}\Gamma_2 \\ \vdots \\ \sqrt{\lambda_k}\Gamma_{\mathrm{k}} \end{bmatrix}$$

$$\hat{\boldsymbol{L}}\hat{\boldsymbol{L}}^{\boldsymbol{T}} = \begin{bmatrix} \hat{l}_{11} & \hat{l}_{12} & \cdots & \hat{l}_{1k} \\ \hat{l}_{21} & \hat{l}_{22} & \cdots & \hat{l}_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{l}_{p1} & \hat{l}_{p2} & \cdots & \hat{l}_{pk} \end{bmatrix} = \left[\ \hat{l}_{i1}, \hat{l}_{i2}, \dots, \hat{l}_{ik}\right] \qquad , i = 1,2, \dots, k$$

Dimana $\mathbf{\Gamma_i}$ adalah kolom ke-i pada matriks $\mathbf{\Gamma}$. Jadi $\hat{\boldsymbol{L}} = (\hat{l}_{ij})$ merupakan penduga matriks bobot faktor **L**. Matriks diagonal ragam khusus $\mathbf{\Psi}$ diduga dengan $\widehat{\mathbf{\Psi}}$, yaitu matriks diagonal yang unsurnya diambil dari $\boldsymbol{R} - \hat{\boldsymbol{L}}\hat{\boldsymbol{L}}'$ dan ditulis sebagai :

$$\widehat{\mathbf{\Psi}} = \begin{bmatrix} 1 - {h_1}^2 & 0 & \cdots & 0 \\ 0 & 1 - {h_2}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 - {h_p}^2 \end{bmatrix}$$

Dimana,

$${h^2}_i = \sum_{i=1}^{k} l_{ij}^2 \qquad , i = 1,2, \dots , p$$

Dengan demikian kita memiliki model *k*-faktor dengan **L** diduga oleh $\hat{\boldsymbol{L}}$ dan $\mathbf{\Psi}$ diduga dengan $\widehat{\mathbf{\Psi}}$**,** sehingga diperoleh pendekatan bagi **R** adalah :

$$\mathbf{R} \approx \hat{\mathbf{L}}\hat{\mathbf{L}}^{\mathbf{T}} + \widehat{\mathbf{\Psi}}$$

Dalam prakteknya, kita harus melakukan validasi model sehingga sebelum menerima $\hat{\boldsymbol{L}}$ dan $\widehat{\mathbf{\Psi}}$ sebagai penduga akhir, perlu dihitung matriks sisaan :

$$\mathbf{Res} = \mathbf{R} - (\hat{\mathbf{L}}\hat{\mathbf{L}}^{\mathbf{T}} + \widehat{\mathbf{\Psi}})$$

Matriks sisaan **Res** selalu memiliki unsur diagonal nol. Pada kasus ideal, **Res = 0**. Dengan demikian jika unsur diagonal juga dekat dengan nilai 0, maka penduga $\hat{\boldsymbol{L}}$ dan $\widehat{\mathbf{\Psi}}$ dianggap cukup bagus dan dapat diterima.

(Mattjik & Sumertajaya, 2011)

2.4.2 Metode Faktor Utama

Metode ini didasarkan pada penggunaan metode komponen utama terhadap matriks korelasi semu (atau matriks ragam-peragam semu), yaitu diperoleh dengan mengganti unsur diagonal dengan komunalitas.

Misalkan struktur faktor $\mathbf{R} = \hat{\mathbf{L}}\hat{\mathbf{L}}^{\mathbf{T}} + \mathbf{\Psi}$ berlaku untuk suatu nilai *k*, seperti sebelumnya *k* adalah banyaknya faktor. Maka matriks korelasi semu $\mathbf{R} - \mathbf{\Psi} = \hat{\mathbf{L}}\hat{\mathbf{L}}^{\mathbf{T}}$ merupakan matriks simetrik semidefinit positif dengan pangkat *k*. Tujuan kita adalah menduga **L**. Menggunakan penguraian spektral terhadap $\mathbf{R} - \mathbf{\Psi}$ menghasilkan

$$\mathbf{R} - \mathbf{\Psi} = \mathbf{\Gamma}\,\mathbf{\Lambda}\,\mathbf{\Gamma}^{\mathbf{T}}$$

dimana $\mathbf{\Lambda}$ adalah diag$(\lambda_1, \dots, \lambda_k, 0, \dots, 0)$ dan $\lambda_1 \geq \cdots \geq \lambda_k > 0$ adalah akar ciri dari matriks korelasi semu dan $\mathbf{\Gamma}$ adalah matriks vektor ciri padanannya. Dengan demikian matriks **L** dapat diperoleh melalui :

$$\hat{\boldsymbol{L}} = \left[\sqrt{\lambda_1}\Gamma_1 | \dots | \sqrt{\lambda_k}\Gamma_{\mathrm{k}}\right]$$

Dalam prakteknya untuk matriks **R** yang diketahui dan $\mathbf{\Psi} = \widehat{\mathbf{\Psi}}$, $\mathbf{R} - \widehat{\mathbf{\Psi}}$ mungkin saja berpangkat penuh. Jadi dengan menggunakan $\widehat{\mathbf{\Gamma}}$ sebagai penduga $\mathbf{\Gamma}$ dan $\widehat{\mathbf{\Lambda}}$ sebagai penduga $\mathbf{\Lambda}$, melalui penguraian

$$\mathbf{R_a} = \mathbf{R} - \mathbf{\Psi} = \widehat{\mathbf{\Gamma}}\,\widehat{\mathbf{\Lambda}}\,\widehat{\mathbf{\Gamma}}^{\mathbf{T}}$$

Kita menduga **L** sebagai

$$\hat{\boldsymbol{L}} = \left[\sqrt{\hat{\lambda}_1}\hat{\Gamma}_1 | \dots | \sqrt{\hat{\lambda}_k}\hat{\Gamma}_{\mathrm{k}}\right]$$

Pendugaan **L** yang dinyatakan diatas tergantung pada $\widehat{\boldsymbol{\Psi}}$ yang ditentukan atau pada komunalitas $\hat{h}_i^2$. Dengan demikian, apabila ingin menggunakan metode ini, maka harus menentukan komunalitasnya. Ada beberapa cara untuk menentukan hal ini, yaitu dengan pendugaan menggunakan *Squared Multiple Correlation* (*SMC)*, pendugaan dengan nilai mutlak korelasi yang terbesar, dan pendugaan *Adjusted Squared Multiple Correlation (ASMC).*

(Mattjik & Sumertajaya, 2011)

2.4.3 Metode Kemungkinan Maksimum

Metode ini mengasumskan bahwa matriks ragam peragam atau matriks korelai semua peubah bersifat non-singular. Karena metode ini bersifat *scale-invariant* maka kita dapa memperoleh solusi menggunakan salah satu matriks tersebut. Misalkan $\boldsymbol{x_1}, \dots, \boldsymbol{x_n}$ adalah contoh acak dari populasi yang menyebar normal ganda-*p* dengan vektor rataan **μ** dan matriks ragam peragam Σ , yang diasumsikan memiliki struktur peragam $\boldsymbol{\Sigma} = \mathbf{LL^T} + \boldsymbol{\psi}$. Permasalahan analisis faktor dapat dipandang sebagai pencarian penduga kemungkinan maksimum bagi **L** dan **ψ.**

Diketahui bahwa rataan vektor contoh $\bar{\mathbf{X}}$ merupakan penduga kemungkinan maksimum bagi **μ.** Untuk mendapatkan penduga kemungkinan maksimum bagi **L** dan **ψ** kita gunakan matriks ragam peragam contoh **S** sebagai penduga awal bagi **Σ.** Karena contoh diasumsikan berasal dari populasi yang menyebar normal ganda, maka (n – 1)**S** memiliki sebaran Wishart dengan derajat bebas (n – 1) dan nilai harapan **S** adalah **Σ**. Fungsi kepekatan peluang bagi **S** adalah L(**S**) diberikan oleh :

$$\boldsymbol{L(S)} = \boldsymbol{c}.\,|\boldsymbol{\Sigma}|^{-\frac{n-1}{2}}\ |\boldsymbol{S}|^{\frac{n-1}{2}-\frac{p+1}{2}}\ \boldsymbol{e}^{\frac{n-1}{2}tr(\boldsymbol{\Sigma}^{-1}\mathbf{S})}$$

Dengan c adalah konstanta. Sehingga log-likelihood dari **L** dan **ψ,** jika $\boldsymbol{\Sigma} = \mathbf{LL^T} + \boldsymbol{\psi}$ adalah

$$\ln c - \left(\frac{n-1}{2}\right) \boldsymbol{tr}\,[\,(\mathbf{LL^T} + \boldsymbol{\psi})^{-1}\,\mathbf{S}\,] - \ln|\,(\mathbf{LL^T} + \boldsymbol{\psi})^{-1}\,\mathbf{S}\,| \qquad (\mathbf{2.9})$$

Penduga kemungkinan maksimum bagi **L** dan **ψ** diperoleh dengan memaksimumkan fungsi (2.9) dengan kendala *k (k* – 1*)*/2 persyaratan kenunikan (Johnson & Wichern, 1998).

Metode yang dapat digunakan untuk menentukan banyaknya faktor bersama, yaitu uji nisbah kemungkinan (*likelihood ratio test*) dan kriteria informasi Akaike.

2.4.4 Metode Kuadrat Terkecil Tak – Terboboti (*Unweighted Least Square Method, ULS)*

Joreskog (1977) mengembangkan algoritma untuk menduga **L** dan **ψ** secara iteratif, yaitu meminimumkan fungsi kuadrat terkecil

$$F_u\,(\mathbf{L}\,,\boldsymbol{\psi}) = {}^{1}\!/_{2}\ \mathrm{tr}\,(\mathbf{S} - \boldsymbol{\Sigma})^2 = {}^{1}\!/_{2}\ \mathrm{tr}\,[(\mathbf{S} - \boldsymbol{\Sigma})(\mathbf{S} - \boldsymbol{\Sigma})]$$

Terhadap **L** dan **ψ,** dimana $\boldsymbol{\Sigma} = \mathbf{LL^T} + \boldsymbol{\psi}$, dan $\mathbf{LL^T}$ adalah matriks diagonal berukuran *k x k.* Solusi permasalahan ini diperoleh dengan menggunakan algoritma yang disebut ULS. Meskipun metode ini tidak *scale-invariant*, tapi metode ini memiliki kelebihan karena tidak mengasumsikan mariks ragam

peragam (atau matriks korelasi) harus non-singular. Secara intinya, algoritma untuk metode ULS ini sama dengan algoritma iteratif dua tahap.

2.5 Pencilan

Dalam Hampel et al (1986), *outlier* atau pencilan didefinisikan sebagai observasi yang terletak sangat jauh dari bagian terbesar data dan membahayakan beberapa prosedur statistik klasik. Dengan kata lain, pencilan merupakan penyimpangan beberapa observasi dalam suatu data yang terletak cukup jauh dari pusat observasi. Data pencilan adalah pengamatan yang dapat diidentifikasikan secara jelas yang berbeda dari pengamatan yang lain, namun data *outlier* dapat menunjukkan karakteristik dari populasi sehingga tidak harus diabaikan atau dibuang. Pemeriksaan atau identifikasi pencilan pada data harus dilakukan karena pencilan memberikan pengaruh pada ragam dan setelah pencilan teridentifikasi maka dapat diputuskan untuk mempertahankan atau menghapus pencilan tersebut (Hair *et al.*, 1992).

Secara umum, identifikasi pencilan didasarkan pada *Mahalanobis Distance* (MD) dan didefinisikan sebagai berikut :

$$MD_i = \sqrt{(x_i - \bar{x})^T S^{-1} (x_i - \bar{x})} \qquad \textbf{(2.10)}$$

dimana,

x_i = data objek ke-i

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

S = matriks *variance covariance*

untuk $i = 1, \dots, n$, x_i merupakan *outlier* jika $MD > X^2_{\alpha,p}$. Prinsip MD adalah menghitung jarak diruang multidimensional antara sebuah pengamatan dengan pusat dari semua pengamatan (Hair *et al.*, 1998).

Selain menggunakan metode MD, terdapat juga cara lain untuk mengatasi adanya pencilan yaitu dengan menerapkan metode pendugaan yang tegar terhadap pencilan yang dikenal dengan metode *robust* (Johnson D.E, 1998). Suatu pencilan akan lebih teridentifikasi dengan jelas ketika menggunakan metode robust dibandingkan dengan menggunakan metode klasik (*mahalanobis distance*). Dengan menggantikan nilai tengah dan matriks kovarian dari persamaan umum MD dengan nilai tengah dan matriks kovarian robust maka jarak mahalanobis (*mahalanobis distance*) disebut sebagai jarak robust (Filzmoser, 2004).

2.6 Robust

Suatu prosedur statistika dikatakan robust apabila tidak begitu peka terhadap penyimpangan asumsi-asumsi yang mendasarinya (Kendall & Buckland, 1981). Prosedur robust ditujukan untuk mengakomodasi adanya keanehan data, sekaligus meniadakan identifikasi adanya data pencilan dan juga bersifat otomatis dalam menanggulangi data pencilan.

Untuk mengatasi masalah pencilan ini Rousseeuw memperkenalkan metode robust yang dapat bekerja dengan baik meskipun dalam keadaan data yang mengandung pencilan, yaitu metode *Minimum Volume Ellipsoid* (MVE). Cara kerja metode ini adalah mencari suatu ellipsoid yang memuat h pengamatan dari keseluruhan data yang berukuran n yang memiliki volume paling minimum.

MVE digunakan untuk jumlah pengamatan yang lebih besar dari jumlah variabel (Rousseeuw & Zomeren, 1990, 1991).

2.7 Metode *Minimum Volume Ellipsoid (MVE)*

Minimum Volume Ellipsoid (MVE) yang pertama kali diperkenalkan oleh Rousseuw merupakan penaksir robust dengan *high breakdown* pertama pada lokasi *scatter* multivariat yang sering digunakan. MVE terkenal karena kepekaannya terhadap pencilan yang membuatnya dapat diandalkan untuk mendeteksi pencilan dan tersedia secara luas, serta implementasi MVE mudah dioperasikan pada algoritma perhitungannya (Rousseuw & Van Aelst, 2009).

Prinsip utama dari metode MVE adalah membuat atau menentukan sebuah elipsoida yang volumenya minimum dari sekumpulan data, sedemikian sehingga titik-titik data yang berada diluar elips dikatakan sebagai pencilan (*outlier*). Jika n adalah banyaknya pengamatan dan p adalah banyaknya variabel bebas, maka terdapat sebanyak kombinasi $p+1$ dari n atau $\binom{n}{p+1}$ himpunan bagian (sub sampel) yang memuat $p+1$ pengamatan. Dari data dalam jumlah besar inilah, banyaknya elips yang harus dibuat, lalu ditentukan elips yang paling minimum volumenya. Dengan kata lain, metode ini berfungsi untuk mencari elipsoid dengan volume paling minimum yang melingkupi suatu sub himpunan dari minimal h pengamatan pada himpunan $X = \{x_1, .., x_n\} \in \mathbb{R}^p$, dengan penaksir lokasi MVE merupakan pusat elipsoid dan penaksir *scatter* MVE berkoresponden menjadi matriks bentukan (*shape matrix).*

MVE didefinisikan sebagai pasangan $(\bar{\boldsymbol{X}}, \boldsymbol{S})$ dengan penaksir lokasi $\bar{\boldsymbol{X}}$ adalah vektor-p dan penaksir *scatter* **S** adalah matriks semi-definit positif pxp sedeikian sehingga determinan **S** diminimalkan untuk memenuhi**,**

$$\{(\boldsymbol{x_i} - \bar{\boldsymbol{X}})' \, \boldsymbol{S^{-1}} \, (\, \boldsymbol{x_i} - \, \bar{\boldsymbol{X}}) \leq \boldsymbol{c^2}\} \geq \boldsymbol{h}$$

dimana n adalah jumlah observasi, p adalah jumlah variabel, $h = [\frac{n+p+1}{2}]$, c adalah konstanta dan $\bar{\boldsymbol{X}}$ adalah himpunan data pengamatan (Rousseuw dan Van Aelst, 1990,1991).

Menurut Rousseuw dan Van Aelst (2009), jarak robust $RD(x_i)$ yang berhubungan dengan MVE didefinisikan:

$$RD(x_i) = \sqrt{(\boldsymbol{x_i} - \bar{\boldsymbol{X}}) \, \boldsymbol{S^{-1}} \, (\boldsymbol{x_i} - \bar{\boldsymbol{X}})^{\boldsymbol{t}}}, \qquad i = 1, \dots, n. \qquad \textbf{(2.11)}$$

Karena dalam mencari solusi eksak dalam menyelesaikan masalah pada MVE ini dapat dikatakan sulit, maka alternatif lain adalah melalui pendekatan algoritma resampling yang disebut juga algoritma (p + 1)-subset. Algoritma ini memiliki keuntungan dalam penaksirannya karena masih bersifat *affine equivariant* dan masih mempertahankan *high breakdown point*.

Breakdown point adalah kelompok terkecil adanya pencilan yang mengakibatkan suatu penaksir menghasilkan penaksiran yang jauh berbeda atau bias. Konsep *breakdown* dilakukan untuk mengetahui kemampuan suatu penaksir dalam menghasilkan suatu penaksir dalam menghasilkan nilai taksiran yang resisten terhadap adanya pencilan dalam jumlah tertentu (Akbar & Maftukhah, 2007).

2.8 Algoritma Resampling

Algoritma resampling didefinisikan sebagai proses mengambil sejumlah subsampel secara acak, yang memuat p+1 pengamatan. Untuk setiap subsampel, dihitung vektor nilai tengah dan matriks kovariannya. Ellipsoid dengan volume terkecil selanjutnya digunakan untuk menduga **μ** dan **Σ (**Suhupi, 2006).

Misalkan diberikan data hasil *n* observasi (pengamatan) dengan *p* variabel acak sebagai berikut :

Pengamatan	X_1	X_2	$\cdots$	X_p
1	x_{11}	x_{12}	$\cdots$	x_{13}
2	x_{21}	x_{22}	$\cdots$	x_{23}
3	x_{31}	x_{32}	$\cdots$	x_{33}
4	x_{41}	x_{42}	$\cdots$	x_{43}
5	x_{51}	x_{52}	$\cdots$	x_{53}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
n	$x_{\mathrm{n}1}$	$x_{\mathrm{n}2}$	$\cdots$	$x_{\mathrm{n}3}$

Dari data *n* pengamatan ini, akan terdapat subsampel sebanyak $\binom{n}{p+1}$ yang memuat (p+1) pengamatan.

Untuk mendapatkan elipsoida dengan volume minimum, dilakukan langkah-langkah sebagai berikut :

1. Bentuk subsampel yang mengandung p+1 pengamatan yang disimbolkan dengan indeks $J = \{i_1, i_2, \dots, i_{p+1}\} \subset \{1,2, \dots, n\}$
2. Menentukan rata-rata (mean) sampel dengan rumus :

$$\boldsymbol{T_J} = \overline{\boldsymbol{x}}_{\boldsymbol{J}} = \frac{1}{p+1} \sum_{i=1}^{p+1} x_{iJ} \tag{2.12}$$

Selanjutnya ditentukan matriks kovarian dari sampel dengan rumus :

$$\boldsymbol{S_J} = \frac{1}{p} \sum_{i=1}^{p+1} (\boldsymbol{x_{iJ}} - \overline{\boldsymbol{x}}_{\boldsymbol{J}})(\boldsymbol{x_{iJ}} - \overline{\boldsymbol{x}}_{\boldsymbol{J}})^{\ T} \qquad (2.13)$$

Matriks kovarian $\boldsymbol{S_J}$ ini pada umumnya non singular dan memang harus non singular. Bila tidak, banyaknya pengamatan harus ditambah sampai pemilihan (p+1) himpunan bagian dari sampel ini mengakibatkan $\boldsymbol{S_J}$ non singular.

3. Menghitung jarak Mahalanobis untuk semua pengamatan dengan rumus :

$$D_J^2 = \left[(\boldsymbol{x_{iJ}} - \overline{\boldsymbol{x}}_{\boldsymbol{J}})\ \ \boldsymbol{S_J^{-1}}\ \ (\boldsymbol{x_{iJ}} - \overline{\boldsymbol{x}}_{\boldsymbol{J}})^{\ T}\right]_{h:n} \qquad (2.14)$$

dimana *h:n* menunjukkan jarak kuadrat terkecil ke-*h* diantara jarak kuadrat pada *n* pengamatan dalam *X*. Dengan menggunakan jarak kuadrat tersebut, dibuat faktor skala $\frac{D_J^2}{c^2}$ dimana $c = \sqrt{X_{p,\propto}^2}$. Nilai $\frac{D_J}{c}$ ini merupakan jari-jari elips ke arah sumbu X_J.

4. Sebagai penentu yang sangat penting dalam MVE adalah volume ellipsoida yang poposional dengan nilai :

$$V_J = \left(\frac{D_J}{c}\right)^p \sqrt{\det(\boldsymbol{S_J})} \qquad (2.15)$$

5. Setelah selesai langkah ke-4 dengan memperoleh nilai volume ellips untuk subsampel pertama, maka ulangi langkah ke 1 sampai 4 diatas untuk subsampel ke 2 berukuran sama yaitu (p + 1) pengamatan sehingga diperoleh volume elips ke 2. Proses diulang sampai sebanyak $\binom{n}{p+1}$ subsampel.

6. Pilih subsampel yang elipsnya memiliki volume paling minimum.

7. Hitung $\boldsymbol{T}(\boldsymbol{X}) = \boldsymbol{T_J}\ \ dan\ \boldsymbol{S}(\boldsymbol{X}) = \frac{c^2(n,p)}{X_{p:\alpha}^2}\ D_J^2\ S_J$ dari V_J yang akan digunakan sebagai estimasi mean dan matriks varians kovarians, dengan $c^2(n,p) =$

$\left[1+\frac{15}{n-p}\right]^2$ yang disebut dengan *correction term*. Sedangkan $\boldsymbol{\chi^2_{p,\alpha}}$ menyatakan faktor koreksi untuk menentukan konsistensi karena data berasal dari sebaran normal multivariat dan memiliki titik *breakdown* maksimum sebesar 50%.

8. Berdasarkan T(X) dan S(X), akan dihitung jarak robust dengan rumus:

$$RD_i = \sqrt{(\boldsymbol{x_i} - \boldsymbol{T}(\boldsymbol{X}))\ \ \boldsymbol{S}(\boldsymbol{x})^{-1}\ \ (\boldsymbol{x_i} - \boldsymbol{T}(\boldsymbol{X}))^{\ T}}$$

9. Selanjutnya pencilan data ditentukan apabila :

$$RD_i > \sqrt{\chi^2_{P,\alpha}}$$

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester ganjil tahun akademik 2017/2018 dan bertempat di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung.

3.2 Data Penelitian

Data yang digunakan dalam penelitian ini adalah data yang dibangkitkan dengan simulasi pada *software SAS*, yaitu data berukuran n = 100 dan p = 5 variabel dari sebaran normal multivariat serta tentukan terlebih dahulu *v*ektor *mean* dan matriks *variance-covariance*. Berdasarkan perhitungan, untuk membangkitkan data pada *SAS* digunakan vektor *mean* dan matriks *variance-covariance* sebagai berikut :

$$\boldsymbol{\mu} = \begin{bmatrix} 4 \\ 2 \\ 3 \\ 2 \\ 5 \end{bmatrix}, \qquad \boldsymbol{\Sigma} = \begin{bmatrix} 19.3 & -3.25 & 0.75 & -3.25 & -0.5 \\ -2.25 & 6 & 0.75 & 0.75 & 2.75 \\ 0.75 & 0.75 & 13 & -1 & -3 \\ -3.25 & 0.75 & -1 & 14 & 1.25 \\ -0.5 & 2.75 & -3 & 1.25 & 16 \end{bmatrix} \tag{3.1}$$

Pada penelitian ini disimulasikan suatu kondisi dimana data tersebut akan diberi pencilan sebanyak 0%, 5%, 10%, 15%, 20%, 25% dan 30%.

3.3 Metode Penelitian

Dalam penelitian ini terutama pada penulisan skripsi dilakukan dengan menggunakan studi literatur secara sistematis yang bersumber dari buku-buku atau media lain untuk memperoleh informasi sebanyak mungkin guna mendukung penulisan skripsi. Untuk mempermudah perhitungan digunakan *software SAS.*

Berikut langkah-langkah yang dilakukan dalam penelitian ini adalah :

1. Membangkitkan data yang berukuran n = 100 dan p = 5 variabel dari sebaran normal multivariat dengan vektor *mean* dan matriks *variance-covariance* pada persamaan (3.1). Data yang diperoleh dimisalkan data $\mathbf{X}^{(0)}$.
2. Menghitung matriks korelasi dan *robust* dari data $\mathbf{X}^{(0)}$**.** Matriks korelasi klasik diperoleh dengan menggunakan persamaan (2.4), sedangkan matriks korelasi *robust* diperoleh menggunakan metode MVE dengan algoritma resampling.
3. Menduga bobot faktor $\mathbf{L}^{(0)}$ dan ragam spesifik $\boldsymbol{\psi}^{(0)}$ berdasarkan matriks korelasi klasik dan *robust*. Kemudian, bobot faktor dikalikan dengan transposenya $\boldsymbol{A}^{(0)} = \boldsymbol{L}^{(0)}\left(\boldsymbol{L}^{(0)}\right)^{\boldsymbol{t}}$.
4. Membangkitkan sebuah matriks noise ($\boldsymbol{E}^{(s)}$) dari sebaran normal (0, $(0.05)^2$).
5. Membangkitkan matriks pencilan ($\boldsymbol{P}^{(s)}$). Matriks ($\boldsymbol{P}^{(s)}$) disini elemennya adalah nol kecuali untuk beberapa elemen yang dijadikan pencilan.
6. Menambahkan matriks ($\boldsymbol{E}^{(s)}$) dan ($\boldsymbol{P}^{(s)}$) pada data $\mathbf{X}^{(0)}$**.** Sehingga diperoleh matriks $\mathbf{X}^{(0)}$ yang telah terkontaminasi :

$$\boldsymbol{X}^{(S)} = \boldsymbol{X}^{(0)} + \boldsymbol{E}^{(s)} + \boldsymbol{P}^{(s)} \quad , untuk\ s = 1, \dots, m;\ \ m = 100$$

7. Menduga matriks korelasi klasik dan *robust* dari data $\boldsymbol{X}^{(\boldsymbol{S})}$

8. Menduga bobot faktor $\boldsymbol{L}^{(\boldsymbol{s})}$ dan ragam spesifik $\boldsymbol{\psi}^{(\boldsymbol{s})}$ berdasarkan matriks korelasi klasik dan *robust*.

9. Mengalikan bobot faktor $(\boldsymbol{L}^{(\boldsymbol{s})})$ dengan matriks transposenya $\boldsymbol{A}^{(\boldsymbol{s})} = \boldsymbol{L}^{(\boldsymbol{s})}\left(\boldsymbol{L}^{(\boldsymbol{s})}\right)^{\boldsymbol{T}}$.

10. Ulangi langkah 4 sampai 9 sebanyak 100 kali

11. Membandingkan elemen-elemen $a^{(s)}$ dari $\boldsymbol{A}^{(\boldsymbol{s})} = \boldsymbol{L}^{(\boldsymbol{s})}\left(\boldsymbol{L}^{(\boldsymbol{s})}\right)^{\boldsymbol{T}}$ dengan $a^{(0)}$ dari matriks $\boldsymbol{A}^{(\boldsymbol{0})} = \boldsymbol{L}^{(\boldsymbol{0})}\left(\boldsymbol{L}^{(\boldsymbol{0})}\right)^{\boldsymbol{t}}$ dan juga untuk ragam spesifiknya $\boldsymbol{\psi}^{(\boldsymbol{s})}$ dengan $\boldsymbol{\psi}^{(\boldsymbol{0})}$**,** dari kedua metode dengan menghitung nilai *mean square error* (MSE) atau nilai kuadrat tengah galat :

$$MSE\left(a_{ij}\right) = \frac{1}{m}\sum_{x=1}^{m}(a_{ij}^{(S)} - a_{ij}^{(0)})^2 \tag{4.1}$$

$$MSE\left(\psi_j\right) = \frac{1}{m}\sum_{x=1}^{m}(\sqrt{P_j^{(S)}} - \sqrt{P_j^{(0)}})^2 \tag{4.2}$$

12. Menghitung rata-rata MSE (a_{ij}) bobot faktor dan MSE (ψ_j) ragam spesifik, yaitu :

$$MSE(A) = \frac{1}{p^2}\sum_{i=1}^{p}\sum_{j=1}^{p} MSE(a_{ij}) \tag{4.3}$$

$$MSE(\psi) = \frac{1}{p}\sum_{j=1}^{p} MSE\left(\psi_j\right) \tag{4.4}$$

13. Membandingkan MSE dari kedua metode yaitu metode klasik dan *robust.*

IV. HASIL DAN PEMBAHASAN

4.1 Hasil Simulasi Analisis Faktor Klasik dan *Robust*.

Hasil dari langkah-langkah simulasi yang telah dilakukan menggunakan *software SAS* untuk data bangkitan pertama serta *mean* dan matriks *variance covariance* pada persamaan 3.1 dengan pencilan 0%, 10%, 15%, 20%, 25% dan 30% adalah sebagai berikut :

1. Membangkitkan data dari sebaran normal multivariat dengan nilai vektor *mean* dan matriks *variance-covariance* pada persamaan 3.1. Hasil *output* data dapat dilihat pada lampiran.
2. Menentukan matriks korelasi klasik dan *robust* dari data $\mathbf{X}^{(0)}$ yang telah dibangkitkan.

   Matriks korelasi klasik :

$$r_k^{(0)} = \begin{bmatrix} 1 & -0.278134 & -0.014251 & -0.169571 & -0.062215 \\ -0.278134 & 1 & 0.0921196 & 0.1239315 & 0.3301533 \\ -0.014251 & 0.0921196 & 1 & -0.10643 & -0.095723 \\ -0.169571 & 0.1239315 & -0.10643 & 1 & 0.3283372 \\ -0.062215 & 0.3301533 & -0.095723 & 0.3283372 & 1 \end{bmatrix}$$

   Mariks korelasi *robust* :

$$r_b^{(0)} = \begin{bmatrix} 1 & -0.278102 & -0.009446 & -0.165058 & -0.06528 \\ -0.278102 & 1 & 0.0920028 & 0.1274889 & 0.3314461 \\ -0.009446 & 0.0920028 & 1 & -0.148745 & -0.088081 \\ -0.165058 & 0.1274889 & -0.148745 & 1 & 0.365205 \\ -0.06528 & 0.3314461 & -0.088081 & 0.365205 & 1 \end{bmatrix}$$

3. Menguraikan matriks korelasi diatas dengan analisis faktor menggunakan metode komponen utama, sehingga diperoleh dugaan bobot faktor dan ragam spesifik seperti dibawah ini :

   Matriks bobot faktor :

$$L_k^{(0)} = \begin{bmatrix} -0.515623 & 0.3838581 \\ 0.6818265 & -0.415074 \\ -0.100234 & -0.791282 \\ 0.6341626 & 0.3576627 \\ 0.7193037 & 0.2430194 \end{bmatrix}, \qquad L_r^{(0)} = \begin{bmatrix} -0.495113 & 0.395839 \\ 0.6603546 & -0.46309 \\ -0.143182 & -0.776183 \\ 0.6645193 & 0.3820474 \\ 0.7373204 & 0.1855038 \end{bmatrix}$$

   Bobot faktor yang telah diperoleh dari data $\mathbf{X^{(0)}}$ dikalikan dengan matriks transposenya $\boldsymbol{A}^{(\mathbf{0})} = \boldsymbol{L}^{(\mathbf{0})}\left(\boldsymbol{L}^{(\mathbf{0})}\right)^{\boldsymbol{t}}$ sehingga diperoleh,

   Bobot faktor klasik :

$$A_k^{(0)} = \begin{bmatrix} 0.4132144 & -0.510895 & -0.252057 & -0.189697 & -0.277605 \\ -0.510895 & 0.6371734 & 0.260098 & 0.2839326 & 0.3895694 \\ -0.252057 & 0.260098 & 0.6361734 & -0.346576 & -0.264395 \\ -0.189697 & 0.2839326 & -0.346576 & 0.5300849 & 0.5430745 \\ -0.277605 & 0.3895694 & -0.264395 & 0.5430745 & 0.5764563 \end{bmatrix}$$

   Bobot faktor *robust* :

$$A_r^{(0)} = \begin{bmatrix} 0.4018258 & -0.510259 & -0.236353 & -0.177783 & -0.291628 \\ -0.510259 & 0.6505205 & 0.2648921 & 0.2618961 & 0.400988 \\ -0.236353 & 0.2648921 & 0.6229617 & -0.391686 & -0.249556 \\ -0.177783 & 0.2618961 & -0.391686 & 0.5875461 & 0.5608349 \\ -0.291628 & 0.400988 & -0.249556 & 0.5608349 & 0.5780531 \end{bmatrix}$$

   Matriks ragam spesifik klasik:

$$\psi_k^{(0)} = \begin{bmatrix} 0.5867856 & 0 & 0 & 0 & 0 \\ 0 & 0.3628266 & 0 & 0 & 0 \\ 0 & 0 & 0.3638266 & 0 & 0 \\ 0 & 0 & 0 & 0.4699151 & 0 \\ 0 & 0 & 0 & 0 & 0.4235437 \end{bmatrix}$$

Matriks ragam spesifik *robust*:

$$\psi_r^{(0)} = \begin{bmatrix} 0.5981742 & 0 & 0 & 0 & 0 \\ 0 & 0.3494795 & 0 & 0 & 0 \\ 0 & 0 & 0.3770383 & 0 & 0 \\ 0 & 0 & 0 & 0.4124539 & 0 \\ 0 & 0 & 0 & 0 & 0.4219469 \end{bmatrix}$$

4. Membangkitkan matriks noise ($\boldsymbol{E}^{(s)}$) yang dibangkitkan dari sebaran normal multivariat dengan n = 100 dan p = 5 variabel.
5. Menambahkan pencilan yang telah ditentukan pada data $\mathbf{X}^{(0)}$ sehingga akan diperoleh matriks pencilan ($\boldsymbol{P}^{(s)}$).
6. Dengan menjumlahkan $\boldsymbol{X}^{(0)}, \boldsymbol{E}^{(s)}$ dan $\boldsymbol{P}^{(s)}$, maka diperoleh matriks $\mathbf{X}^{(s)}$ yang telah terkontaminasi (*s* menyatakan simulasi, *s* = 1,2,...100)
7. Menduga kembali matriks korelasi klasik dan robust pada data $\mathbf{X}^{(s)}$.

   Matriks korelasi klasik :

$$r_k^{(s)} = \begin{bmatrix} 1 & -0.276341 & -0.015426 & -0.167406 & -0.061586 \\ -0.276341 & 1 & 0.0933832 & 0.1272646 & 0.3297683 \\ -0.015426 & 0.0933832 & 1 & -0.105547 & -0.094648 \\ -0.167406 & 0.1272646 & -0.105547 & 1 & 0.3263936 \\ -0.061586 & 0.3297683 & -0.094648 & 0.3263936 & 1 \end{bmatrix}$$

   Mariks korelasi *robust* :

$$r_b^{(s)} = \begin{bmatrix} 1 & -0.276313 & -0.010635 & -0.162929 & -0.064575 \\ -0.276313 & 1 & 0.093305 & 0.1309963 & 0.331021 \\ -0.010635 & 0.093305 & 1 & -0.148115 & -0.087012 \\ -0.162929 & 0.1309963 & -0.148115 & 1 & 0.3627233 \\ -0.064575 & 0.331021 & -0.087012 & 0.3627233 & 1 \end{bmatrix}$$

8. Selanjutnya, bobot faktor yang diperoleh dari data $\mathbf{X}^{(s)}$ dikalikan dengan matriks transposenya $\boldsymbol{A}^{(s)} = \boldsymbol{L}^{(s)}\left(\boldsymbol{L}^{(s)}\right)^{\boldsymbol{t}}$. Hasil kali bobot faktor serta matriks ragam spesifik dari analisis faktor klasik dan *robust* adalah sebagai berikut :

Bobot faktor klasik :

$$A_k^{(s)} = \begin{bmatrix} 0.4111122 & -0.508648 & -0.255867 & -0.188435 & -0.274516 \\ -0.508648 & 0.6360412 & 0.2598207 & 0.2876562 & 0.3910159 \\ -0.255867 & 0.259820 & 0.6387338 & -0.34332 & -0.263183 \\ -0.188435 & 0.2876562 & -0.34332 & 0.5288225 & 0.5427623 \\ -0.274516 & 0.3910159 & -0.263183 & 0.5427623 & 0.5761977 \end{bmatrix}$$

Bobot faktor *robust* :

$$A_r^{(s)} = \begin{bmatrix} 0.399556 & -0.508078 & -0.23981 & -0.176413 & -0.288922 \\ -0.508078 & 0.6494268 & 0.2647469 & 0.265599 & 0.4025153 \\ -0.23981 & 0.2647469 & 0.6261553 & -0.389232 & -0.247908 \\ -0.176413 & 0.265599 & -0.389232 & 0.5862376 & 0.5601433 \\ -0.288922 & 0.4025153 & -0.247908 & 0.5601433 & 0.5770237 \end{bmatrix}$$

Matriks ragam spesifik klasik:

$$\psi_k^{(s)} = \begin{bmatrix} 0.5888878 & 0 & 0 & 0 & 0 \\ 0 & 0.3639588 & 0 & 0 & 0 \\ 0 & 0 & 0.3612662 & 0 & 0 \\ 0 & 0 & 0 & 0.4711775 & 0 \\ 0 & 0 & 0 & 0 & 0.4238023 \end{bmatrix}$$

Matriks ragam spesifik *robust*:

$$\psi_r^{(s)} = \begin{bmatrix} 0.600444 & 0 & 0 & 0 & 0 \\ 0 & 0.3505732 & 0 & 0 & 0 \\ 0 & 0 & 0.3738447 & 0 & 0 \\ 0 & 0 & 0 & 0.4137624 & 0 \\ 0 & 0 & 0 & 0 & 0.4229763 \end{bmatrix}$$

9. Menghitung rata-rata nilai $MSE(a_{ij})$ klasik dan *robust* serta $MSE(\psi_j)$ klasik dan *robust* untuk data dengan seluruh presentase pencilan yang telah ditentukan. Sehingga, diperoleh hasil sebagai berikut :

Tabel 1. Perbandingan Nilai MSE Bobot Faktor serta Ragam Spesifik Klasik dan *Robust* untuk seluruh Presentase Pencilan.

Presentase Pencilan	MSE Bobot Faktor		MSE Ragam Spesifik	
	Klasik	*Robust*	Klasik	*Robust*
0 %	0.0000165836	0.064831798	0.0000138051	0.000821009
5 %	0.495394046	0.181071312	0.191300191	0.005960234
10 %	0.496934899	0.136492835	0.196138531	0.006047493
15 %	0.497317953	0.05058603	0.197761044	0.006729121
20 %	0.49678568	0.022935018	0.198574668	0.015283413
25 %	0.496608836	0.026492442	0.199036294	0.016815358
30 %	0.496780491	0.035556305	0.199295918	0.012742192

Berdasarkan Tabel 1 dapat dilihat untuk data dengan presentase pencilan 0% nilai MSE bobot faktor maupun ragam spesifik yang dihasilkan metode klasik lebih kecil dibandingkan nilai MSE bobot faktor dan ragam spesifik dengan metode *robust*, yaitu 0.0000165836 < 0.064831798 untuk MSE bobot faktor dan 0.0000138051 < 0.000821009 untuk MSE ragam spesifik.

Sedangkan, nilai MSE bobot faktor dan ragam spesifik metode *robust* pada data dengan pencilan sebanyak 5%, 10%, 15%, 20%, 25% dan 30% menghasilkan nilai yang selalu lebih kecil dibandingkan dengan metode klasik. Nilai MSE bobot faktor metode *robust* yang dihasilkan untuk data berpencilan antara lain, 0.181071312, 0.136492835, 0.05058603, 0.022935018, 0.026492442, dan 0.035556305 lebih kecil dari nilai MSE bobot faktor metode klasik, yaitu 0.495394046, 0.496934899, 0.497317953, 0.49678568, 0.496608836, dan 0.496780491. Sama halnya dengan nilai MSE ragam spesifik metode *robust*, yaitu 0.005960234, 0.006047493, 0.006729121, 0.015283413, 0.016815358, dan

0.012742192 jauh lebih kecil dari nilai MSE ragam spesifik klasik, yaitu 0.191300191, 0.196138531, 0.197761044, 0.198574668, 0.199036294 dan 0.199295918.

4.2 Hasil Analisis Data Bangkitan dengan Nilai *Mean* dan Matriks *Variance Covariance* Random

Hasil yang sama juga diperlihatkan oleh data bangkitan lain dengan *mean* dan matriks *variance covariance* berbeda yang dipilih secara random pada *software SAS*. Nilai MSE yang dihasilkan oleh data bangkitan lainnya dirangkum dalam tabel berikut ini :

Tabel 2. Perbandingan Nilai MSE Bobot Faktor serta Ragam Spesifik Klasik dan *Robust* untuk seluruh Presentase Pencilan dan 5 Gugus Data Bangkitan.

Presentase Pencilan	Data	MSE Bobot Faktor		MSE Ragam Spesifik	
		Klasik	*Robust*	Klasik	*Robust*
0 %	1	0.000016185	0.002862399	0.000005531	0.00026996
	2	0.000024643	0.002701663	0.000011024	0.000624214
	3	0.000030551	0.01217522	0.000008208	0.000191726
	4	0.000026968	0.0008726	0.00001039	0.000301159
	5	0.000034098	0.002085346	0.000008672	0.000211039
	Rata-rata	**0.00002648**	**0.004139446**	**0.000008765**	**0.00031962**
5 %	1	0.166850072	0.002846424	0.003963776	0.000150731
	2	0.212600292	0.00262111	0.042759576	0.000890031
	3	0.23562951	0.0118381	0.030234903	0.000193384
	4	0.219079811	0.001582266	0.03182859	0.000220639
	5	0.223409367	0.00201566	0.015531027	0.000290858
	Rata-rata	**0.21151381**	**0.004180712**	**0.024863574**	**0.000349129**

Presentase Pencilan	Data	MSE Bobot Faktor		MSE Ragam Spesifik	
		Klasik	*Robust*	Klasik	*Robust*
10 %	1	0.165889477	0.002022068	0.003971349	0.000112279
	2	0.209591733	0.00254285	0.042845443	0.000790094
	3	0.232498906	0.001408087	0.0302912	0.000094659
	4	0.215322686	0.002592799	0.031881185	0.000162108
	5	0.227832418	0.001812432	0.015556545	0.000339869
	Rata-rata	**0.210227044**	**0.002075647**	**0.024909144**	**0.000299802**
15 %	1	0.166547032	0.001577905	0.003973797	0.000060289
	2	0.207949038	0.001802706	0.042873522	0.000436172
	3	0.231154379	0.006198924	0.030308793	0.000213733
	4	0.213668512	0.003198643	0.031897939	0.000297262
	5	0.210298748	0.003155558	0.015564957	0.000290678
	Rata-rata	**0.210298748**	**0.003155558**	**0.024923802**	**0.000259627**
20 %	1	0.165759655	0.001719134	0.003974975	0.00007777
	2	0.207031919	0.001759196	0.042886439	0.000279561
	3	0.230403673	0.000451067	0.030318035	0.000273895
	4	0.212715429	0.003763667	0.031906399	0.000287858
	5	0.231164075	0.002916799	0.015569177	0.000087979
	Rata-rata	**0.20941495**	**0.002121973**	**0.024931005**	**0.000201413**
25 %	1	0.165255333	0.001492405	0.003975681	0.000061378
	2	0.206433201	0.002305491	0.042894852	0.000091422
	3	0.229924034	0.000510369	0.030323263	0.000443359
	4	0.212110354	0.001853982	0.031911193	0.000111236
	5	0.230524073	0.002091933	0.015571596	0.000215831
	Rata-rata	**0.208849399**	**0.001650836**	**0.024935317**	**0.000184645**
30 %	1	0.164943562	0.004843465	0.003976148	0.000032797
	2	0.206037778	0.007610189	0.042899858	0.000333374
	3	0.229592933	0.006665986	0.030326364	0.001332483
	4	0.211641306	0.00203208	0.031914479	0.000055119
	5	0.230096321	0.004134199	0.015572999	0.000105479
	Rata-rata	**0.20846238**	**0.005057184**	**0.02493797**	**0.00037185**

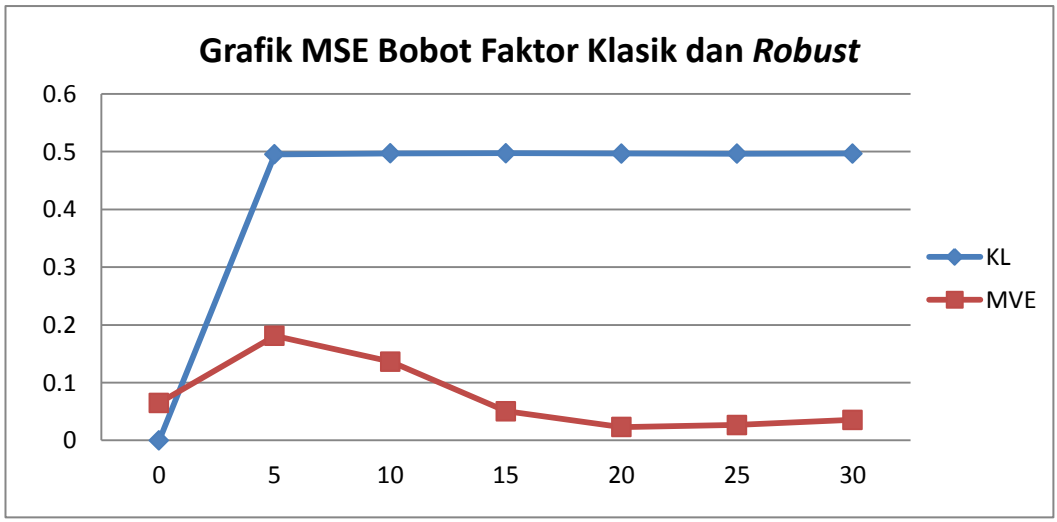
Pada Tabel 2. dapat dilihat bahwa nilai MSE metode klasik untuk seluruh data yang dibangkitkan dengan pencilan 0% menghasilkan nilai yang lebih kecil dari nilai MSE metode *robust.* Sedangkan, untuk data dengan pencilan 5%, 10%, 15%, 20%, 25% dan 30% metode *robust* menghasilkan nilai yang jauh lebih kecil dibandingkan metode klasik.

4.3 Grafik Perbandingan Nilai *Mean Square Error* Analisis Faktor Klasik dan *Robust.*

Nilai MSE pada Tabel 1 dan Tabel 2 ditampilkan juga dalam bentuk grafik. Berikut ini adalah grafik perbandingan nilai MSE bobot faktor metode klasik dan *robust*, serta ragam spesifik klasik dan *robust* :

4.3.1 Gambar Grafik Nilai MSE Tabel 1.

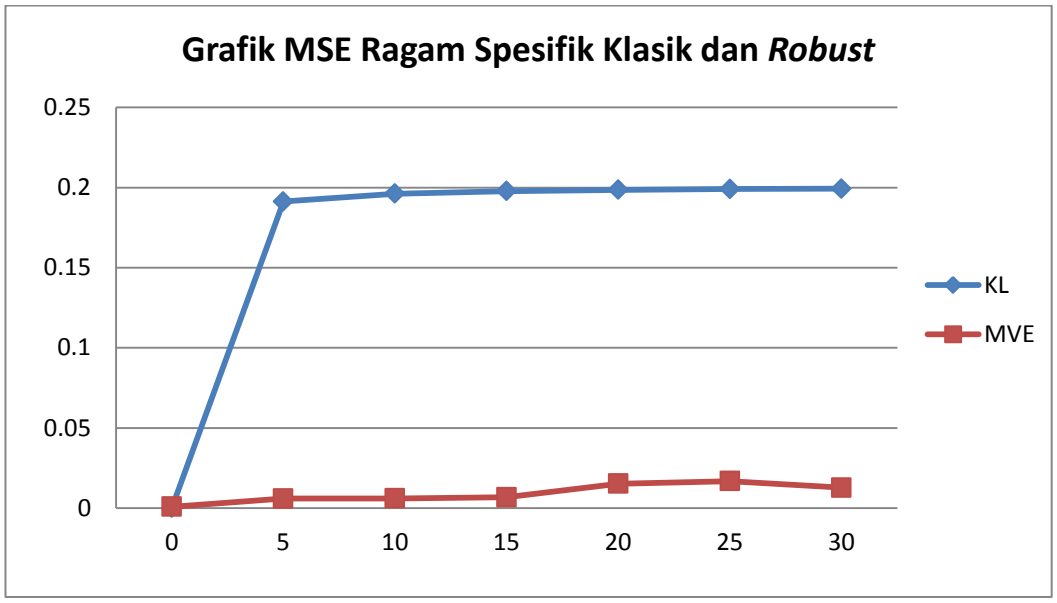
Grafik nilai MSE pada Tabel 1 dari analisis dengan nilai *mean* dan matriks *variance covariance* pada persamaan 3.1 dapat dilihat dibawah ini :



Gambar 1. Grafik Nilai MSE Bobot Faktor Klasik dan *Robust* pada Tabel 1.

Berdasarkan grafik pada Gambar 1, garis berwarna biru menunjukkan nilai MSE ragam spesifik untuk metode klasik, yaitu 0.0000138051 dimana nilai ini lebih rendah dari nilai MSE pada metode *robust*, yaitu 0.064831798 yang ditunjukkan garis berwarna merah pada data dengan pencilan 0%.

Sedangkan, untuk data dengan pencilan 5%, 10%, 15%, 20%, 25% dan 30% dapat dilihat nilai MSE untuk metode klasik meningkat drastis dibandingkan nilai MSE metode *robust* yang cenderung bernilai kecil (nilai dapat dilihat pada Tabel 1.)



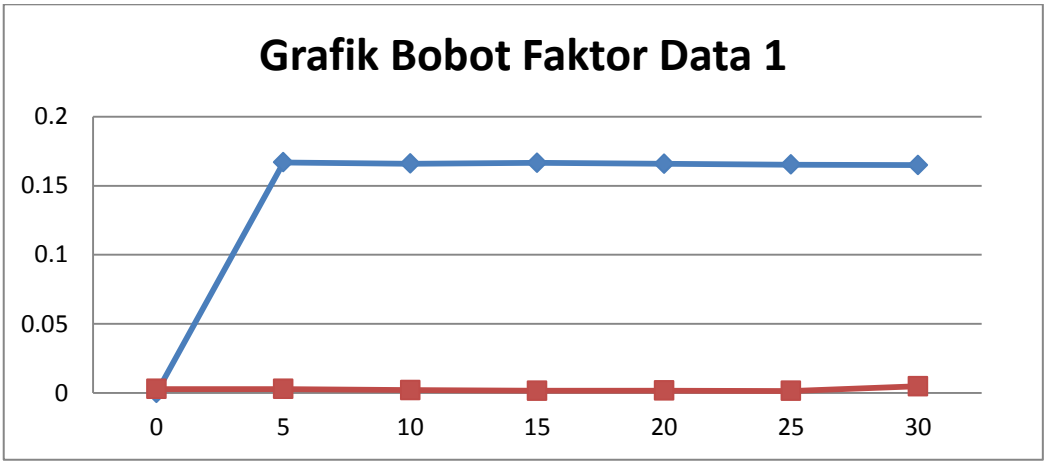
Gambar 2. Grafik Nilai MSE Ragam Spesifik Klasik dan *Robust* pada Tabel 1.

Berdasarkan grafik pada Gambar 2, garis berwarna biru menunjukkan nilai MSE ragam spesifik untuk metode klasik, yaitu 0.0000165836 dimana nilai ini lebih rendah dari nilai MSE pada metode *robust*, yaitu 0.064831798 yang ditunjukkan garis berwarna merah pada data dengan pencilan 0%.

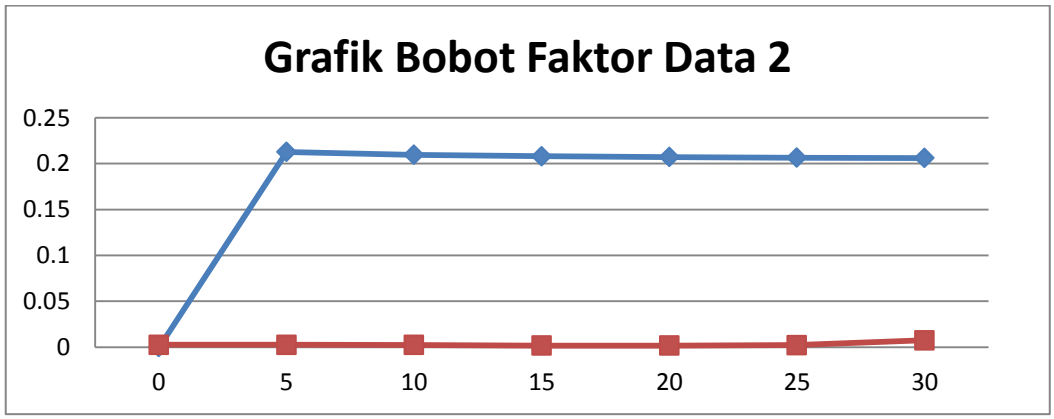
Sedangkan, untuk data dengan pencilan 5%, 10%, 15%, 20%, 25% dan 30% dapat dilihat grafik nilai MSE untuk metode klasik meningkat drastis dibandingkan nilai MSE metode *robust* yang cenderung bernilai lebih kecil dan konstan.

4.3.2 Gambar Grafik Nilai MSE Tabel 2.

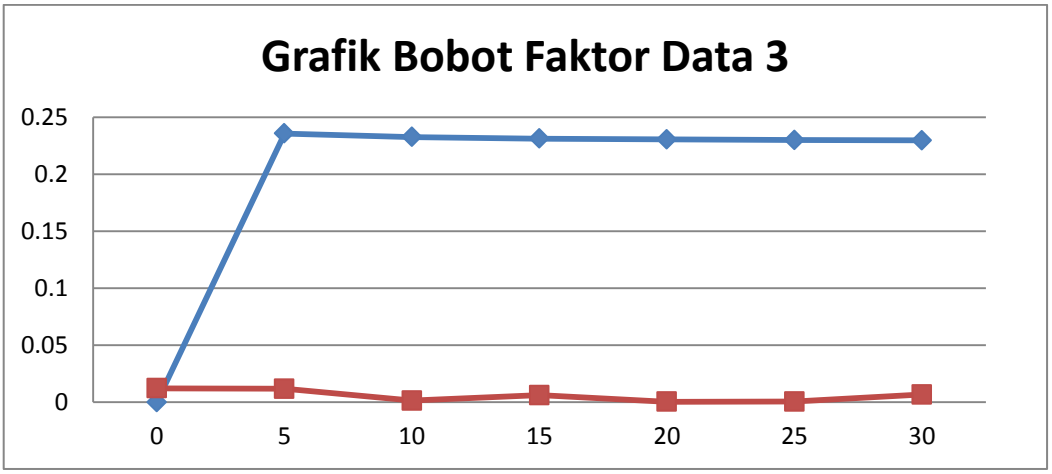
Grafik nilai MSE bobot faktor dan ragam spesifik dari analisis dengan nilai *mean* dan matriks *variance covariance* random untuk seluruh data bangkitan dapat dilihat dibawah ini :



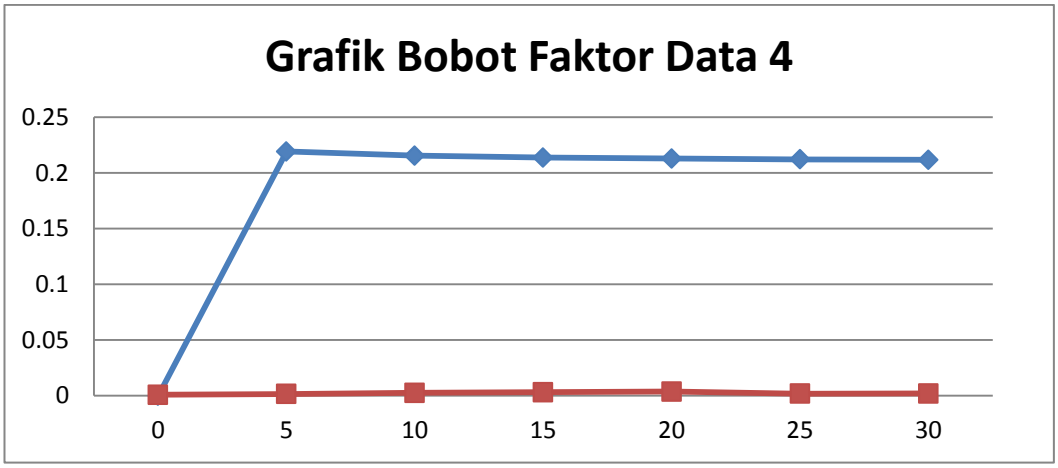
Gambar 3. Grafik Nilai MSE Bobot Faktor Klasik dan *Robust* untuk Data Pertama



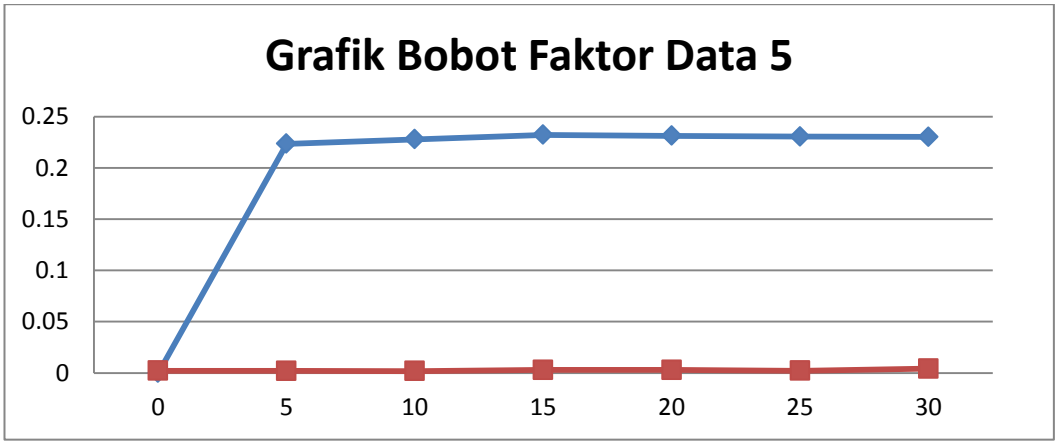
Gambar 4. Grafik Nilai MSE Bobot Faktor Klasik dan *Robust* untuk Data Kedua



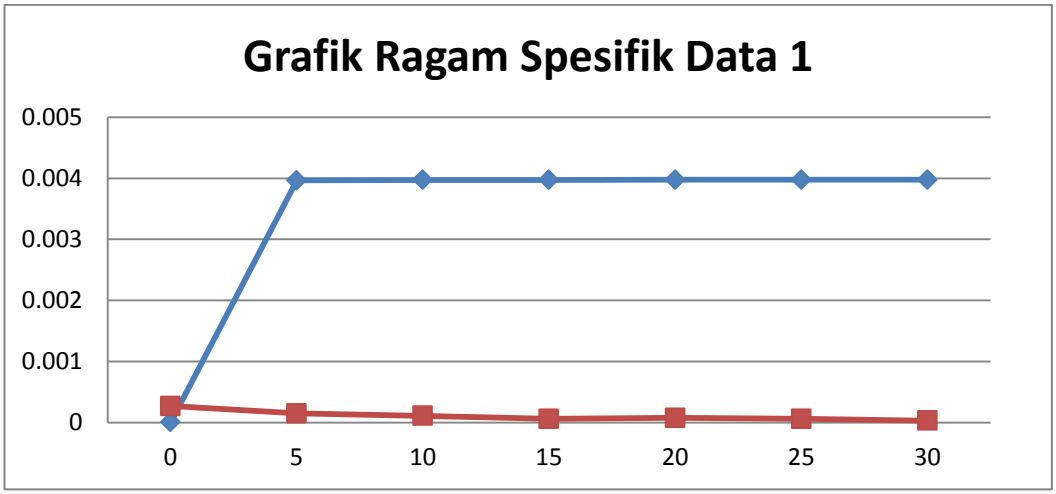
Gambar 5. Grafik Nilai MSE Bobot Faktor Klasik dan *Robust* untuk Data Ketiga



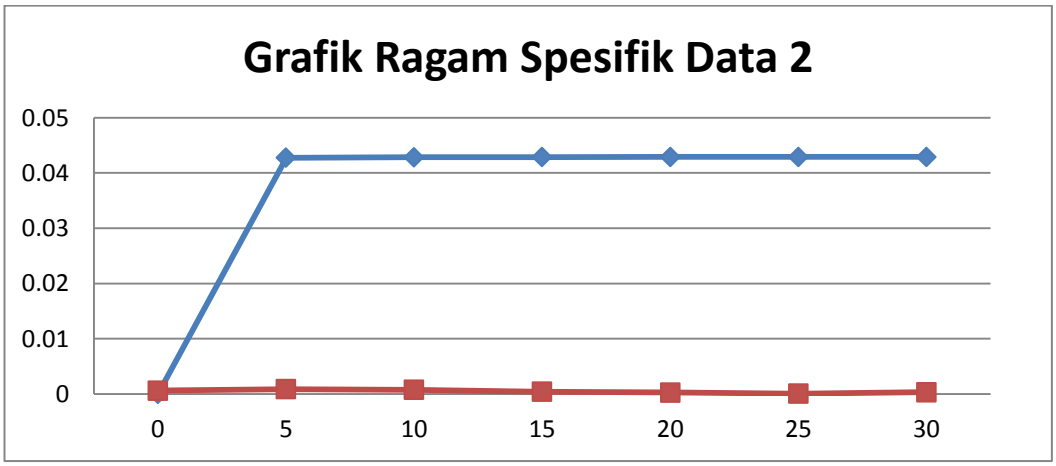
Gambar 6.Grafik Nilai MSE Bobot Faktor Klasik dan *Robust* untuk Data Keempat



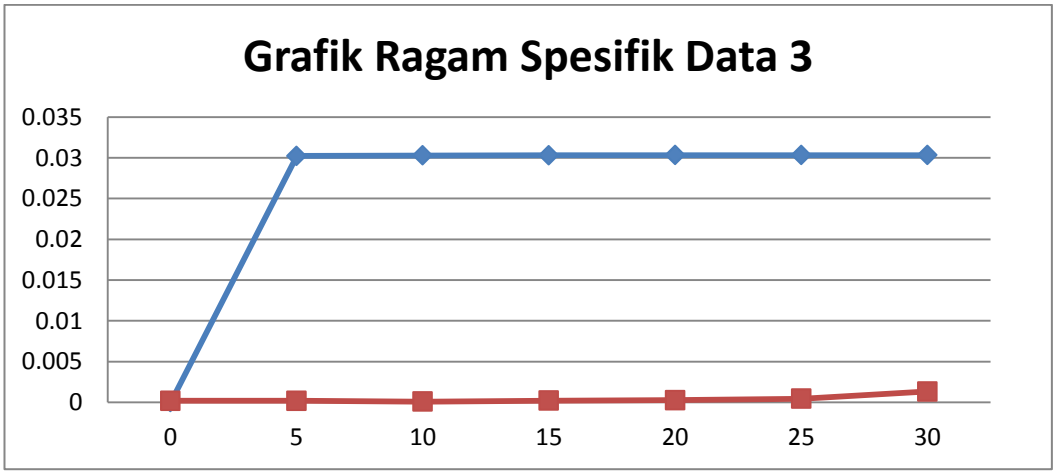
Gambar 7. Grafik Nilai MSE Bobot Faktor Klasik dan *Robust* untuk Data Kelima



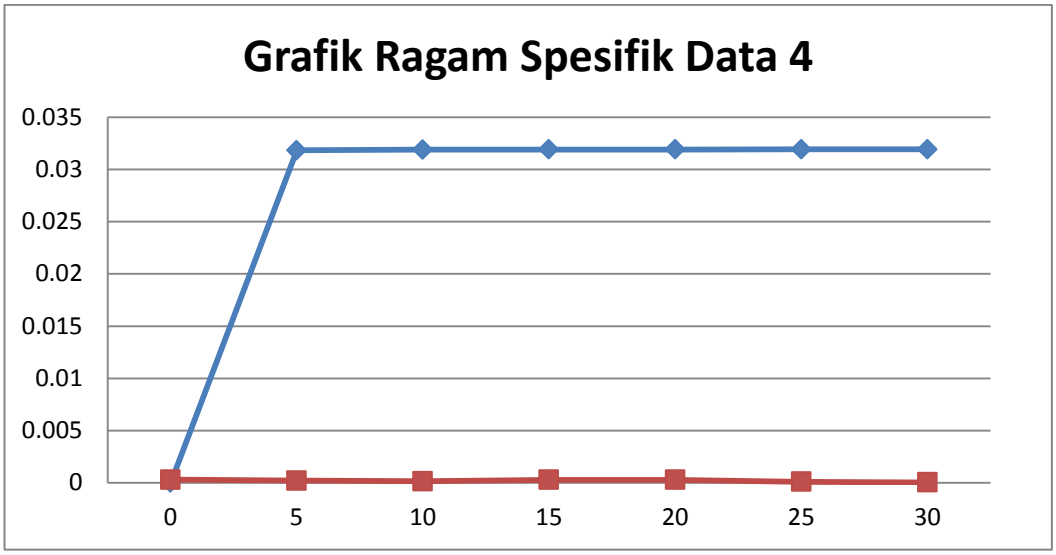
Gambar 8. Grafik Nilai MSE Ragam Spesifik Klasik dan *Robust* Data Pertama



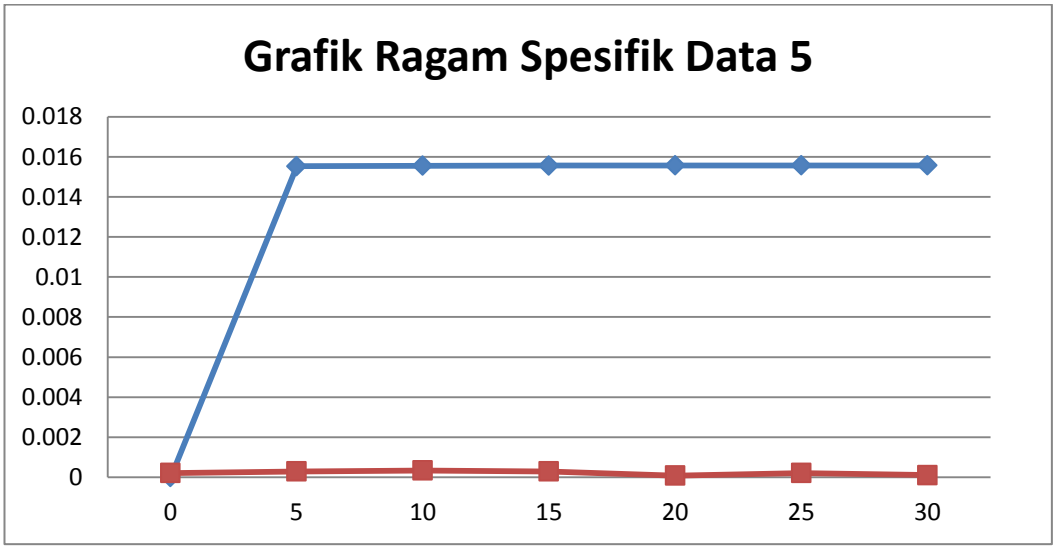
Gambar 9. Grafik Nilai MSE Ragam Spesifik Klasik dan *Robust* Data Kedua



Gambar 10. Grafik Nilai MSE Ragam Spesifik Klasik dan *Robust* Data Ketiga



Gambar 11. Grafik Nilai MSE Ragam Spesifik Klasik dan *Robust* Data Keempat



Gambar 12. Grafik Nilai MSE Ragam Spesifik Klasik dan *Robust* Data Kelima

Berdasarkan grafik-grafik bobot faktor maupun ragam spesifik diatas yang diperoleh dari seluruh data bangkitan, dapat dilihat bahwa perbandingan antara nilai MSE metode klasik dan *robust* sangat jauh berbeda terutama untuk data yang mengandung pencilan. Pada grafik, ditunjukkan untuk data dengan pencilan 0% menghasilkan nilai yang lebih kecil untuk metode klasik (nilai dapat dilihat pada

Tabel 2.) dan meningkat drastis pada saat data mengandung pencilan. Sedangkan, untuk metode *robust* nilai MSE jauh lebih kecil dan konstan pada saat data mengandung pencilan 5%, 10%, 15%, 20%, 25% dan 30% dibandingkan dengan metode klasik.

4.4 Pembahasan

Berdasarkan analisis yang telah dilakukan, diperoleh hasil yang dapat dilihat pada Tabel 1 dan 2 serta gambar grafik-grafiknya. Tabel dan grafik tersebut memperlihatkan perbandingan nilai MSE yang diperoleh untuk analisis faktor klasik dan analisis faktor *robust* pada kasus dimana data mengandung pencilan 0%, 5%, 10%, 15%, 20%, 25% dan 30%.

Data dengan pencilan 0% pada Tabel 1 maupun 2, memperlihatkan nilai MSE bobot faktor dan ragam spesifik yang diperoleh analisis faktor klasik lebih kecil dibandingkan nilai MSE analisis faktor *robust.* Hal ini juga diperlihatkan pada Gambar 1 hingga Gambar 12 yang menunjukkan bahwa grafik MSE klasik (garis berwarna biru) berada dibawah grafik MSE *robust* (garis berwarna merah), yang artinya nilai MSE klasik lebih kecil dari nilai MSE *robust.* Perbandingan tersebut terlihat jelas pada Gambar 1.

Sedangkan, pada saat data mengandung pencilan 5%, 10%, 15%, 20%, 25% dan 30% menunjukkan bahwa nilai MSE analisis faktor *robust* selalu lebih kecil dibandingkan analisis faktor klasik. Hal ini juga dibuktikan pada Gambar 1 hingga 12, dimana grafik nilai MSE bobot faktor maupun ragam spesifik analisis faktor klasik (garis berwarna biru) cenderung meningkat ketika data mengandung

pencilan, sedangkan grafik *robust* (garis berwarna merah) selalu berada di bawah grafik klasik dimana nilainya jauh lebih kecil dan cenderung stabil terhadap perubahan presentase yang diberikan.

Semakin kecil nilai MSE suatu estimator, maka hasil estimasinya akan semakin baik. Sehingga dalam penelitian ini, analisis faktor *robust* dengan metode MVE lebih baik dari analisis faktor klasik dalam mengatasi data berpencilan. Sedangkan untuk data tidak berpencilan, analisis faktor klasik lebih baik dibandingkan analisis faktor *robust*.

V. KESIMPULAN

Dari hasil analisis dan pembahasan maka diperoleh kesimpulan sebagai berikut :

1. Metode *Minimum Volume Ellipsoid* (MVE) pada analisis faktor *robust* dengan titik awal kerja yaitu matriks korelasi, digunakaan untuk menduga nilai MSE bobot faktor dan ragam spesifik. Kedua parameter tersebut merupakan aspek penting dalam analisis faktor dan nilai MSE digunakan sebagai acuan untuk melihat bahwa analisis faktor *robust* baik digunakan dalam mengatasi data berpencilan.
2. Nilai MSE yang dihasilkan analisis faktor *robust* selalu lebih kecil dari analisis faktor klasik pada kondisi dimana data mengandung pencilan. Sedangkan, pada kondisi data dengan pencilan 0% maka nilai MSE analisis faktor klasik akan jauh lebih kecil dibandingkan analisis faktor *robust*. Oleh karena itu, analisis faktor klasik baik digunakan untuk data dengan pencilan 0% dan analisis faktor *robust* baik digunakan untuk data pengamatan dengan pencilan 5%, 10%, 15%, 20%, 25% dan 30%.

DAFTAR PUSTAKA

Mattjik, A.A & Sumertajaya, I.M. 2011. *Sidik Peubah Ganda Dengan Menggunakan SAS*. Bogor : IPB PRESS.

Akbar, M. S. & Maftukhah, L. 2007. Optimasi Kekuatan Tourque Pada Lampu TL. *Jurnal Ilmiah Sains dan Teknologi* **6(3):** hal 218 - 229.

Anton, H. & Rorres, C. 2004. *Aljabar Linear Elementer*. Edisi 8. (diterjemahkan oleh :Indriasari, R., dan Harmein, I). Jakarta : Erlangga.

Suhupi, A. 2006. *Pendugaan Matriks Kovarian Robust Dengan Metode MVE Untuk Mendeteksi Pencilan Pada Analisis Multivariat.* Skripsi. Jurusan Matematika, Bandar Lampung.

Filzmoser, P. 2004. A Multivariate Outlier Detection Method. In Aivazian, S., Filzmoser, P. And Kharin, Y. (editors). Proceedings Of The Seventh International On Computer Data Analysis Modelling. 1 : 18-22.

Hair, J.F. JR., Anderson, R.E, Tatham, R.L. & Black, W.C. 1992. *Multivariate Data Analysis with Readings.* New York : Mcmillan Publishing Co.

Hair, J.F. JR., Anderson, R.E, Tatham, R.L. & Black, W.C. 1998. *Multivariate Data Analysis,* Fifth Edition. New Jersy : Prentice Hall, Upper Saddle River .

Hampel F.R., Ronchetti E.M., Rousseeuw P.J., & Stahel W. A. 1986. *Robust Statistics. The Approach Based on Influence Functions*. New York : John Wiley and Sons.

Johnson, D.E. 1998. *Applied Multivariate Methods for Data Analysis.* Duxbury Press, Pacific Grove.

Johnson, R.A. & Wichern, D. W. 2007. *Applied Multivariate Sttistical Analysis.* 6th edition. New Jersey : Printice Hall.

Rousseeuw, P.J. & Van Zomeren, B.C. 1990. Unmasking Multivariate Outlier and Leverage Points. *Journal of the American Statistical Association.* 85: 633-639.

Rousseeuw, P.J. & Van Zomeren, B.C. 1991. Robust Distance : Simulations and Cutoff Values. *The IMA Volumes in Mathematics and its Application.* 34: 195-203.

Suliyanto. 2005. *Analisis Data Dalam Aplikasi Pemasaran.* Bogor : Ghalia Indonesia.