EVALUASI KORELASI KANONIK KLASIK DAN KORELASI KANONIK *ROBUST* PADA DATA INDEKS KESEJAHTERAAN MASYARAKAT TERHADAP TINGKAT PERTUMBUHAN EKONOMI DI INDONESIA

(Skripsi)

Oleh

MUFLIHAH
NPM 1817031002

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2022

ABSTRACT

EVALUATION OF CLASSIC CANONICAL CORRELATION AND ROBUST CANONICAL CORRELATION ON COMMUNITY WELFARE INDEX DATA TO ECONOMIC GROWTH RATE IN INDONESIA

By

MUFLIHAH

Canonical correlation analysis is a multivariate statistical model that focuses on the correlation between linear combinations of two sets of variables, so that the correlation is maximized. Outliers in the data affect the resulting covariance variance matrix, so that the covariance variance matrix is inefficient and the estimator is biased. One of the methods used is to use a robust estimator. The multivariate approach to robust canonical correlation analysis is the Minimum Covariance Determinant (MCD) method. This study aims to determine the evaluation of the results of the correlation with the classical canonical correlation method and the robust canonical correlation on the community welfare index data on the level of economic growth in Indonesia. The results of this study indicate that the robust canonical correlation value better explains the correlation between the two sets of variables, the classical canonical correlation coefficient is obtained with the correlation value $\rho_1 = 0.9397$, $\rho_2 = 0.8143$, $\rho_3 = 0.4660$, dan $\rho_4 = 0.2670$, while the robust canonical correlation coefficient is obtained with correlation values $\rho_1 = 0.9879$, $\rho_2 = 0.8602$, $\rho_3 = 0.7012$, dan $\rho_4 = 0.5768$, classical canonical correlation and robust canonical correlation can be interpreted further because it fulfills the significance test in whole and in part.

Keywords: Canonical Correlation Analysis, Covariance Variance Matrix, Robust Estimator, Minimum Covariance Determinant (MCD).

ABSTRAK

EVALUASI KORELASI KANONIK KLASIK DAN KORELASI KANONIK *ROBUST* PADA DATA INDEKS KESEJAHTERAAN MASYARAKAT TERHADAP TINGKAT PERTUMBUHAN EKONOMI DI INDONESIA

Oleh

MUFLIHAH

Analisis korelasi kanonik adalah model statistika multivariat yang berfokus pada korelasi antara kombinasi linear dari dua himpunan variabel, sehingga korelasinya menjadi maksimum. Pencilan pada data mempengaruhi matriks varian kovarian yang dihasilkan, sehingga matriks varian kovariannya tidak efisien dan sifat penduganya menjadi bias. Salah satu cara yang digunakan adalah dengan menggunakan penduga *robust*. Pendekatan multivariat untuk analisis korelasi kanonik *robust* adalah dengan metode *Minimum Covariance Determinant* (MCD). Penelitian ini bertujuan untuk menentukan evaluasi hasil korelasi dengan metode korelasi kanonik klasik dan korelasi kanonik *robust* pada data indeks kesejahteraan masyarakat terhadap tingkat pertumbuhan ekonomi di Indonesia. Hasil penelitian ini menunjukkan bahwa nilai korelasi kanonik *robust* lebih menjelaskan korelasi diantara kedua himpunan variabel, diperoleh koefisien korelasi kanonik klasik dengan nilai korelasi $\rho_1 = 0.9397$, $\rho_2 = 0.8143$, $\rho_3 = 0.4660$, dan $\rho_4 = 0.2670$, sedangkan koefisien korelasi kanonik *robust* diperoleh dengan nilai korelasi $\rho_1 = 0.9879$, $\rho_2 = 0.8602$, $\rho_3 = 0.7012$, dan $\rho_4 = 0.5768$, korelasi kanonik klasik dan korelasi kanonik *robust* dapat diinterpretasikan lebih lanjut karena memenuhi uji signifikansi secara keseluruhan dan secara sebagian.

Kata Kunci: Analisis Korelasi Kanonik, Matriks Varian Kovarian, Penduga *Robust*, *Minimum Covariance Determinant* (MCD).

EVALUASI KORELASI KANONIK KLASIK DAN KORELASI KANONIK *ROBUST* PADA DATA INDEKS KESEJAHTERAAN MASYARAKAT TERHADAP TINGKAT PERTUMBUHAN EKONOMI DI INDONESIA

Oleh

MUFLIHAH

Skripsi

Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA MATEMATIKA

Pada

Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Lampung

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2022

Judul Skripsi : **EVALUASI KORELASI KANONIK KLASIK DAN KORELASI KANONIK *ROBUST* PADA DATA INDEKS KESEJAHTERAAN MASYARAKAT TERHADAP TINGKAT PERTUMBUHAN EKONOMI DI INDONESIA**

Nama Mahasiswa : ***Muflihah***

Nomor Pokok Mahasiswa : **1817031002**

Jurusan : **Matematika**

Fakultas : **Matematika dan Ilmu Pengetahuan Alam**

MENYETUJUI

1. Komisi Pembimbing

Prof. Drs. Mustofa Usman, M.A., Ph.D.
NIP 19570101 198404 1 001

Dr. Notiragayu, S.Si. M.Si.
NIP 19731109 200012 2 001

2. Ketua Jurusan Matematika

Dr. Aang Nuryaman, S.Si., M.Si.
NIP 19740316 200501 1 001

MENGESAHKAN

1. Tim Penguji

Ketua	: **Prof. Drs. Mustofa Usman, M.A., Ph.D.**	....
Sekretaris	: **Dr. Notiragayu, S.Si. M.Si.**	....
Penguji Bukan Pembimbing	: **Drs. Eri Setiawan, M.Si.**	....

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Dr. Eng. Suripto Dwi Yuwono, S.Si., M.T.
NIP 19740705 200003 1 001

Tanggal Lulus Ujian Skripsi : **07 Juni 2022**

PERNYATAAN SKRIPSI MAHASISWA

Yang bertanda tangan di bawah ini :

Nama Mahasiswa : Muflihah

Nomor Pokok Mahasiswa : 1817031002

Jurusan : Matematika

Judul Skripsi : **EVALUASI KORELASI KANONIK KLASIK DAN KORELASI KANONIK *ROBUST* PADA DATA INDEKS KESEJAHTERAAN MASYARAKAT TERHADAP TINGKAT PERTUMBUHAN EKONOMI DI INDONESIA**

Dengan ini menyatakan bahwa penelitian ini adalah hasil pekerjaan saya sendiri dan apabila kemudian hari terbukti bahwa skripsi ini merupakan salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung,

Muflihah

RIWAYAT HIDUP

Penulis bernama Muflihah lahir di Serang pada 03 Agustus 2000. Penulis merupakan anak pertama dari dua bersaudara dari pasangan Bapak Trio Teguh Muhamadi dan Ibu Neliyana.

Penulis mengawali pendidikan di Taman Kanak-Kanak (TK) Al-Fatah pada tahun 2005-2006. Kemudian menempuh pendidikan Sekolah dasar di SDN Ranjeng pada tahun 2006 – 2012. Kemudian melanjutkan ke Sekolah Menengah Pertama di SMPN 1 Ciruas pada tahun 2012-2015, Sekolah Menengah Atas di SMAN 1 Ciruas pada tahun 2015-2018. Pada tahun 2018 penulis terdaftar sebagai mahasiswa Program Studi S1 Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur SNMPTN.

Selama menjadi mahasiswa penulis aktif dalam mengikuti kegiatan organisasi Himpunan Mahasiswa Matematika (HIMATIKA) FMIPA UNILA sebagai anggota aktif. Pada Tahun 2021 penulis melakukan Kuliah Kerja Praktik (KP) di Badan Pusat Statistik Kabupaten Serang serta mengikuti kegiatan Kuliah Kerja Nyata (KKN) di Desa Cihaseum, Pandeglang-Banten. Penulis mengikuti kegiatan Merdeka Belajar Kampus Merdeka (MBKM) Kampus Mengajar 1 pada semester 6 serta mengikuti kegiatan Studi Independen Bersertifikat PT. Sekolah Integrasi Digital pada semester 7.

KATA INSPIRASI

"Sesungguhnya bersama kesulitan ada kemudahan."

(Q.S Al-Insyirah : 6)

"Jangan engkau bersedih, sesungguhnya Allah bersama kita."

(Q.S At-Taubah: 40)

"Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya."

(Q.S Al-Baqarah: 286)

"Dan hanya kepada Tuhanmulah hendaknya kamu berharap."

(Q.S Al-Insyirah : 8)

"Jangan menjelaskan tentang dirimu kepada siapa pun, karena yang menyukaimu tidak butuh itu. Dan yang membencimu tidak percaya itu."

(Ali bin Abi Thalib)

Setiap orang ada masanya, setiap masa ada waktunya. Janganlah pernah menyerah untuk melakukan sesuatu hal, tetaplah fokus pada tujuan akhir.

(Muflihah)

PERSEMBAHAN

Alhamdulillah, puji dan syukur kepada Allah SWT atas nikmat serta hidayahnya sehingga skripsi ini dapat terselesaikan dengan baik dan tepat pada waktunya.

Oleh karena itu, dengan rasa syukur dan bahagia saya persembahkan rasa terimakasih saya kepada :

Bapak Trio Teguh Muhamadi dan Ibu Neliyana

Terimakasih kepada kedua orang tuaku atas segala doa, pengorbanan, motivasi, dan ridho kalian serta dukungannya selama ini. Terimakasih telah memberikan pelajaran berharga kepada anakmu ini tentang makna perjalanan hidup yang sebenarnya sehingga kelak bisa manjadi orang yang bermanfaat bagi semua orang.

Dosen Pembimbing dan Pembahas

Terimakasih kepada dosen pembimbing dan pembahas yang sudah sangat membantu, memberikan motivasi, memberikan arahan serta ilmu yang berharga.

Sahabat-sahabatku

Terimakasih kepada semua orang-orang baik yang telah memberikan pengalaman, semangat, motivasinya, serta doa-doanya dan senantiasa memberikan dukungan dalam hal apapun

Almamater Tercinta Universitas Lampung

SANWACANA

Segala puji dan syukur penulis ucapkan kepada Allah SWT atas segala nikmat dan karunia-Nya yang tak terhingga sehingga penulis dapat menyelesaikan skripsi yang berjudul "**Evaluasi Korelasi Kanonik Klasik dan Korelasi Kanonik *Robust* pada Data Indeks Kesejahteraan Masyarakat Terhadap Tingkat Pertumbuhan Ekonomi di Indonesia**". Dalam penulisan skripsi ini tidak dapat terselesaikan tanpa adanya bimbingan, bantuan dan dukungan dari berbagai pihak. Sehingga, dalam kesempatan ini penulis mengucapkan terimakasih kepada:

1. Kedua Orang Tuaku, Bapak Trio Teguh Muhamadi dan Ibu Neliyana yang selalu memberikan doa, motivasi serta dukungannya.
2. Bapak Prof. Dr. Mustofa Usman, M.A., Ph.D. selaku dosen pembimbing I yang senantiasa membimbing, memberi masukan, saran serta mendukung penulis dalam menyelesaikan skripsi ini.
3. Ibu Dr. Notiragayu, S.Si. M.Si. selaku dosen pembimbing II yang senantiasa memberikan bimbingan, pengarahan, serta saran sehingga penulis dapat menyelesaikan skripsi ini.
4. Bapak Drs. Eri Setiawan, M.Si. selaku dosen penguji yang telah memberikan kritik dan saran yang membangun sehingga skripsi ini dapat terselesaikan.
5. Ibu Dr. Notiragayu, S.Si. M.Si. selaku dosen pembimbing akademik yang telah memberikan bimbingan dan arahan selama masa perkuliahan.
6. Bapak Dr.Aang Nuryaman, S.Si., M.Si., selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Bapak Dr. Eng. Suripto Dwi Yuwono, S.Si., M.T. selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
8. Seluruh dosen, staff, karyawan Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

9. Keluarga sekalian yang selalu memberikan doa serta semangat kepada penulis
10. Sepupu terbaik saya Teh Dea yang selalu memberikan doa, semangat serta mendengarkan keluh kesah saya selama masa perkuliahan.
11. Teman dekat saya Fajar yang selama ini selalu memberikan doa, motivasi, masukan dan menemani penulis dalam keadaan apapun serta mendengarkan keluh kesah saya selama masa perkuliahan.
12. Teman terbaik saya Ira dan Aninda yang selalu memberikan doa, semangat, merangkul dalam keadaan apapun, serta menjadi teman healing ketika stress.
13. Teman terbaik saya Intan dan Risha yang selalu memberikan doa, semangat, masukan serta menemani penulis dalam keadaan apapun yang selalu sabar.
14. Teman satu kosan saya Maydia, Mazi, Silvi, dan Virda yang senantiasa selalu memberikan semangat, serta selalu berbaik hati ketika direpotkan.
15. Semua teman sejurusan dan kelas A yang telah membantu serta memberikan semangat kepada penulis yang mana tidak bisa disebutkan satu persatu.
16. Orang-orang baik yang namanya tidak bisa saya sebutkan satu persatu yang telah menjadi teman terbaik penulis yang selalu memberikan semangat dan menemani penulis dalam keadaan apapun serta telah memberikan pengalaman dan banyak cerita selama masa perkuliahan.
17. Teman-teman seperbimbingan yang selalu memberikan dukungan dan motivasi serta doa-doanya.
18. Semua pihak yang tidak dapat disebutkan satu persatu yang telah membantu penulis dalam menyelesaikan skripsi ini.

Penulis menyadari bahwa masih banyak kekurangan dalam penulisan skripsi ini. Oleh karena itu, penulis mengharapkan masukan serta saran untuk dijadikan pelajaran kedepannya.

Bandar Lampung, 20 April 2022
Penulis,

Muflihah

DAFTAR ISI

Halaman

DAFTAR TABEL

DAFTAR GAMBAR

Gambar Halaman

I. PENDAHULUAN

1.1 Latar Belakang

Analisis Multivariat dapat didefinisikan sebagai salah satu teknik dalam statistika yang digunakan untuk menganalisis secara bersama variabel lebih dari satu (Wustqa, dkk, 2018). Pada sebagian permasalahan penelitian statistika multivariat, ketertarikan peneliti terkadang tidak hanya sebatas pada pembentukan model regresi linear antara variabel terikat dengan variabel bebas saja, tetapi peneliti tertarik pada hubungan linear antara dua himpunan variabel. Salah satu metode yang digunakan dalam analisis multivariat untuk mengidentifikasi hubungan linear antara dua himpunan variabel yaitu analisis korelasi kanonik (Hair, *et al*., 2009).

Analisis korelasi kanonik berfokus pada korelasi antara kombinasi linear dari himpunan variabel terikat dengan kombinasi linear dari himpunan variabel bebas, sehingga korelasi di antara kedua himpunan variabel tersebut menjadi maksimum (Asbah, dkk, 2013). Kombinasi linear yang terbentuk pada masing-masing himpunan variabel dinamakan variabel kanonik. Analisis korelasi kanonik tidak dilakukan antar pasangan variabel asal, akan tetapi dilakukan antar variabel kanonik pada kedua himpunan (Anderson, 2003).

Matriks varian kovarian pada analisis korelasi kanonik sangat sensitif terhadap adanya pencilan. Analisis korelasi kanonik klasik tidak bisa bekerja dengan baik apabila data yang dianalisis terdapat pencilan. Pencilan adalah pengamatan yang tidak mengikuti sebagian besar pola dan terletak jauh dari pusat data (Barnett and

Lewis, 1978). Pencilan tersebut dapat mengakibatkan sebaran data menjadi tidak normal, sehingga matriks varian kovariannya tidak efisien dan sifat penduganya menjadi bias (Barrera and Yohai, 2006). Salah satu cara yang dapat digunakan agar hasil analisis korelasi kanonik tetap optimal pada data yang mengandung pencilan ialah dengan menggunakan penduga *robust.*

Terdapat beberapa penelitian sebelumnya yang menerapkan korelasi kanonik. Salah satunya adalah penelitian Lestari, dkk (2020) untuk mengetahui hubungan antara faktor multidimensi terhadap derajat kemiskinan di Indonesia dengan menggunakan analisis korelasi kanonik. Selain itu, terdapat pula penelitian Riana, dkk (2015) tentang perbandingan dua metode *robust* yaitu *Biweight Midcovariance* dan *Minimum Covariance Determinant* dalam analisis korelasi kanonik menggunakan data simulasi. Perbedaan penelitian ini dengan yang peneliti lakukan adalah pada metode, data, dan variabel untuk mengetahui keeratan hubungan (korelasi) antara indeks kesejahteraan masyarakat terhadap tingkat pertumbuhan ekonomi berdasarkan provinsi di Indonesia tahun 2020.

Penduga *robust* ialah suatu metode dengan kinerja yang baik dan tidak terpengaruh untuk data yang mengandung pencilan (Rousseeuw and Driessen, 1999). Analisis korelasi kanonik *robust* termasuk suatu pendekatan analisis komponen utama apabila data yang diperoleh mengandung pencilan. Ketika analisis korelasi kanonik klasik kurang tepat digunakan untuk data yang mengandung pencilan, maka diharapkan analisis korelasi kanonik dengan penduga *robust* dapat digunakan untuk mengatasi masalah tersebut.

Terdapat beberapa penduga *robust* yang dapat digunakan dalam mengestimasi matriks varian kovarian pada analisis korelasi kanonik, yaitu *Minimum Covariance Determinant* (MCD), *Biweight Midcovariance*, *Project Pursuit, Alternating Regression,* dan *Sign Test*. Penduga *robust* yang akan digunakan pada penelitian ini adalah *Minimum Covariance Determinant* (MCD). MCD memiliki matriks varian kovarian yang dihasilkan dan menjadi alternatif sebagai pengganti matriks varian kovarian pada korelasi kanonik klasik, serta dapat mendeteksi seluruh pencilan dalam data dan memiliki ketahanan yang cukup besar terhadap pencilan (Rousseeuw and Driessen, 1999).

Pada penelitian ini akan diterapkan data indeks kesejahteraan masyarakat dan tingkat pertumbuhan ekonomi di Indonesia. Pertumbuhan ekonomi dapat dijadikan sebagai tolak ukur kesejahteraan masyarakat. Menurut Avatara (2013), pertumbuhan ekonomi didefinisikan sebagai indikator keberhasilan tercapainya pembangunan ekonomi yang baik yang ditandai oleh meningkatnya kesejahteraan rakyat. Indonesia terus mengusahakan demi terciptanya perekonomian yang baik dan dapat meningkatkan kesejahteraan rakyat. Pertumbuhan ekonomi di Indonesia setiap tahunnya mengalami pertumbuhan yang berfluktuatif. Indikator tingkat pertumbuhan ekonomi pada penelitian ini adalah persentase penduduk miskin, tingkat pengangguran terbuka (TPT), indeks pembangunan manusia (IPM), dan persentase produk domestik regional bruto (PDRB) atas dasar harga berlaku.

Pada dasarnya pencapaian kesejahteraan rakyat dilakukan dengan bermacam perubahan dalam pembangunan rakyat yang bertujuan untuk perbaikan kondisi ekonomi, sosial, dan budaya. Kesejahteraan rakyat adalah kondisi terpenuhnya kebutuhan material, spiritual, dan sosial warga negara agar mampu hidup layak serta meningkatkan diri sehingga dapat melaksanakan fungsi sosialnya. Pembangunan ekonomi diharapkan dapat meningkatkan kesejahteraan masyarakat secara merata. Indikator kesejahteraan masyarakat pada penelitian ini adalah persentase perempuan melahirkan dibantu tenaga medis, persentase penduduk yang memiliki jaminan kesehatan, persentase rumah tangga dengan sumber penerangan listrik, persentase status kepemilikan rumah sendiri, persentase rumah tangga dengan sumber air minum yang layak, dan persentase penduduk.

Berdasarkan permasalahan yang telah dipaparkan maka penulis tertarik untuk meneliti mengenai evaluasi korelasi kanonik klasik dan korelasi kanonik *robust*. Oleh karena itu, penelitian ini berjudul “Evaluasi Korelasi Kanonik Klasik dan

Korelasi Kanonik *Robust* pada Data Indeks Kesejahteraan Masyarakat Terhadap Tingkat Pertumbuhan Ekonomi di Indonesia”.

1.2 Tujuan Penelitian

Adapun tujuan penelitian ini adalah sebagai berikut:

1. Menerapkan analisis korelasi kanonik dengan menggunakan metode *Minimum Covariane Determinant* (MCD) pada data yang mengandung pencilan.
2. Menentukan evaluasi hasil keeratan hubungan (korelasi) dengan metode korelasi kanonik klasik dan korelasi kanonik *robust* pada data indeks kesejahteraan masyarakat terhadap tingkat pertumbuhan ekonomi di Indonesia.

1.3 Manfaat Penelitian

Adapun manfaat yang dapat diambil dari penelitian ini adalah:

1. Memperluas wawasan mengenai analisis korelasi kanonik *robust* dengan menggunakan metode *Minimum Covariance Determinant* (MCD) dalam mengatasi data yang mengandung pencilan.
2. Dapat dijadikan sebagai bahan rujukan dan informasi pada bidang statistika khususnya dalam mengembangkan keilmuan dan pengetahuan tentang analisis korelasi kanonik.
3. Hasil korelasi berdasarkan provinsi di Indonesia dapat memberikan informasi dan masukan dalam menentukan kebijakan-kebijakan selanjutnya bagi pemerintah.

II. TINJAUAN PUSTAKA

2.1 Analisis Multivariat

Analisis Multivariat dapat didefinisikan sebagai salah satu teknik dalam statistika yang digunakan untuk menganalisis secara simultan variabel lebih dari satu (Wustqa, dkk, 2018). Secara umum, sampel data pada analisis multivariat sebagai berikut:

$$\boldsymbol{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1q} \\ x_{21} & x_{22} & \cdots & x_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nq} \end{bmatrix} \quad (2.1)$$

$$\boldsymbol{Y} = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1p} \\ y_{21} & y_{22} & \cdots & y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \dots & y_{np} \end{bmatrix} \quad (2.2)$$

dengan:

x_{np}: pengamatan ke- n pada variabel independen ke-q

y_{nq}: pengamatan ke- n pada variabel dependen ke-p

2.2 Analisis Korelasi Kanonik

Analisis korelasi kanonik ialah model statistika multivariat yang berkaitan dengan kombinasi linear antara dua set variabel (Rencher, 2002). Tujuan dari analisis korelasi kanonik adalah untuk mendapatkan gambaran sederhana tentang struktur korelasi antar himpunan bagian dari variabel (Bilodeau and Brenner, 1999). Kelompok variabel bebas dan variabel terikat saling ketergantungan (*dependency*), sehingga dapat diketahui pengaruh dari satu kelompok terhadap kelompok lainnya. Ciri data untuk analisis korelasi kanonik yaitu semua data yang akan dilakukan untuk analisis korelasi kanonik bertipe metrik (interval atau rasio). Dengan demikian, data bertipe non-metrik (nominal atau ordinal) sebaiknya tidak diproses dengan analisis korelasi kanonik (Nugroho, 2008). Analisis korelasi kanonik berfokus pada korelasi antara kombinasi linear dari himpunan variabel terikat $\boldsymbol{Y}^T = (y_1, y_2, \dots, y_p)$ dengan kombinasi linear dari himpunan variabel bebas $\boldsymbol{X}^T = (x_1, x_2, \dots, x_q)$. Ide dari analisis korelasi kanonik yaitu mendapatkan pasangan dari kombinasi linear yang memiliki korelasi terbesar. Kemudian, mencari pasangan dari kombinasi linear di antara pasangan yang tidak berkorelasi pada pasangan kombinasi linear pertama. (Asbah, dkk, 2013). Pasangan kombinasi linear disebut sebagai variabel kanonik sedangkan hubungan diantara pasangan tersebut disebut korelasi kanonik. Analisis korelasi kanonik berusaha memusatkan hubungan dimensi tinggi antara dua himpunan variabel ke dalam sedikit pasangan variabel kanonik (Johnson and Wichern, 2007).

2.2.1 Asumsi Korelasi Kanonik

Dalam analisis korelasi kanonik terdapat beberapa asumsi yang harus dipenuhi diantaranya setiap variabel berdistribusi normal multivariat, variabel bersifat linear, serta tidak terdapat multikolinearitas (Nugroho, 2008).

a. Normal Multivariat

   Analisis korelasi kanonik mampu mengakomodasi variabel metrik tanpa adanya asumsi normalitas yang ketat. Akan tetapi, normalitas dapat memungkinkan korelasi tertinggi antar variabel dan diperlukan untuk pengujian signifikansi fungsi kanonik. Asumsikan bahwa dua himpunan bagian dari variabel y_1 dan y_2 memiliki distribusi normal bersama,

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \sim N_P \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \right) \tag{2.3}$$

Dengan vektor rata-rata $\boldsymbol{\mu}$ dan matriks varian kovarian S (Bilodeau and Brenner, 1999). Menurut Johnson and Wichern (2007), untuk menguji apakah suatu himpunan data berdistribusi normal multivariat adalah dengan melihat plot *chi-square*. Dengan langkah-langkah sebagai berikut:

1. Menghitung nilai jarak kuadrat dengan rumus:

$$d_j^2 = (x_j - \bar{x})^T \boldsymbol{S}^{-1} (x_j - \bar{x}) \tag{2.4}$$

   dimana,

   $j = 1,2, \dots, n$

   dengan:

   d_j^2 : jarak kuadrat

   $x_1, x_2, \dots, x_n$: observasi sampel

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

   $\boldsymbol{S}$: matriks varian kovarian

2. Mengurutkan jarak kuadrat dari yang terkecil sampai yang terbesar, yaitu $d_1^2 \leq d_2^2 \leq \dots \leq d_n^2$.
3. Menggambar plot tersebut dengan koordinat $\left(d_j^2, \chi_p^2\left(\frac{j-\frac{1}{2}}{n}\right)\right)$, dimana $\chi_p^2\left(\frac{j-\frac{1}{2}}{n}\right)$ adalah $\frac{100\left(j-\frac{1}{2}\right)}{n}$ persentil dari distribusi *chi-square* dengan derajat bebas p.

Data dikatakan berdistribusi normal apabila plot membentuk garis lurus (linear) atau paling tidak 50% dari nilai d_j^2 lebih kecil dari $\chi_{p,0.05}^2$ (Johnson and Wichern, 2007).

b. Linearitas

Asumsi linearitas diperlukan untuk menguji apakah ada hubungan yang bersifat linear antara variabel terikat dengan variabel bebas (Mattjik dan Sumertajaya, 2011). Linearitas dikatakan penting untuk analisis korelasi kanonik dan mempengaruhi dua aspek hasil korelasi kanonik. Pertama, koefisien korelasi kanonik antara sepasang variabel kanonik memiliki hubungan linear. Kedua, analisis korelasi kanonik memaksimumkan hubungan linear antara himpunan variabel. Linearitas dapat di uji dengan cara melihat nilai $p - value$ pada tabel $F - statistics$ dari setiap model yang terbentuk dengan hipotesis sebagai berikut:

- Hipotesis:

  H_0: fungsi tidak linear

  H_1: fungsi linear
- Taraf signifikansi:

$$\alpha = 0.05$$

- Daerah kritis:

  H_0 ditolak jika $p - value < 0.05$
- Statistik uji:

$$p - value < 0.05$$

- Keputusan:

  H_0 ditolak apabila $p - value < 0.05$

c. Non-Multikolinearitas

Multikolinearitas berkaitan dengan situasi dimana terdapat hubungan linear baik yang pasti ataupun mendekati pasti diantara variabel bebas. Pada korelasi kanonik tidak boleh terjadi multikolinearitas antar anggota kelompok variabel bebas (Mattjik dan Sumertajaya, 2011). Dalam mendeteksi tidak adanya multikolinearitas dapat dilakukan dengan melihat nilai *tolerance* dan *variance inflation factor* (VIF), dimana nilai $tolerance < 0.01$ atau sama dengan nilai $VIF < 10$ dengan hipotesis sebagai berikut:

- Hipotesis:

  H_0: variabel independen tidak terjadi multikolinearitas

  H_1: variabel independen terjadi multikolinearitas
- Taraf signifikansi:

$$\alpha = 0.05$$

- Daerah kritis:

  H_0 tidak ditolak jika nilai $Tolerance < 0.01$ atau nilai $VIF < 10$
- Statistik uji:

  Nilai $Tolerance < 0.01$ atau nilai $VIF < 10$

- Keputusan:

  H_0 tidak ditolak apabila nilai $Tolerance < 0.01$ atau nilai $VIF < 10$

2.3 Variabel Kanonik dan Korelasi Kanonik

Analisis korelasi kanonik memusatkan perhatian pada korelasi antara kombinasi linear dari variabel-variabel suatu himpunan dengan variabel-variabel di himpunan yang lain (Johnson and Wichren, 2007). Pasangan kombinasi linear disebut variabel kanonik dan korelasinya disebut korelasi kanonik. Variabel terikat $y = (y_1, y_2, \dots, y_p)$ yang dinotasikan dengan vektor variabel acak $\boldsymbol{Y}$ dan variabel bebas $x = (x_1, x_2, \dots, x_q)$ yang dinotasikan dengan vektor variabel acak $\boldsymbol{X}$, dimana $p \leq q$. Maka karakteristik dari vektor variabel acak $\boldsymbol{X}$ dan $\boldsymbol{Y}$ tersebut dapat ditulis sebagai berikut (Johnson and Wichren, 2007):

$$\begin{gathered} E(\boldsymbol{X}) = \boldsymbol{\mu_X} \qquad Cov(\boldsymbol{X}) = \boldsymbol{S_{XX}} \\ E(\boldsymbol{Y}) = \boldsymbol{\mu_Y} \qquad Cov(\boldsymbol{Y}) = \boldsymbol{S_{YY}} \\ Cov(\boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{S_{XY}} = \boldsymbol{S_{YX}} \end{gathered} \tag{2.5}$$

Maka tiap vektor observasi suatu sampel dipartisi menjadi:

$$\begin{bmatrix} \boldsymbol{Y} \\ \cdots \\ \boldsymbol{X} \end{bmatrix} = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{ip} \\ \cdots \\ x_{i1} \\ x_{i2} \\ \vdots \\ x_{iq} \end{bmatrix}, i = 1,2, \dots, n \tag{2.6}$$

(Johnson and Wichren, 2007).

Untuk tiap sampel pada n vektor observasi, maka vektor rata-rata:

$$\boldsymbol{\mu} = \begin{bmatrix} E(\boldsymbol{X}) \\ E(\boldsymbol{Y}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu_X} \\ \boldsymbol{\mu_Y} \end{bmatrix} \tag{2.7}$$

atau dapat ditulis:

$$\begin{bmatrix} \bar{\boldsymbol{Y}} \\ \cdots \\ \bar{\boldsymbol{X}} \end{bmatrix} = \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \vdots \\ \bar{y}_p \\ \cdots \\ \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_q \end{bmatrix} \tag{2.8}$$

dengan:

$\bar{\boldsymbol{Y}}$: Rata-rata sampel variabel terikat

$\bar{\boldsymbol{X}}$: Rata-rata sampel variabel bebas

$\bar{y}_p$: matriks rata-rata partisi ke-p variabel terikat

$\bar{x}_q$: matriks rata-rata partisi ke-q variabel bebas

(Johnson and Wichren, 2007).

Matriks varian kovarian sampel keseluruhan untuk $y_1, y_2, \dots, y_p$ dan $x_1, x_2, \dots, x_q$ dapat dipartisi sebagai:

$$\boldsymbol{S} = \begin{bmatrix} S_{XX} & S_{XY} \\ S_{YX} & S_{YY} \end{bmatrix} \tag{2.9}$$

dengan:

S_{XX}: matriks varian sampel berukuran $q \times q$ dari $x_1, x_2, \dots, x_q$

S_{YX}: matriks kovarian sampel berukuran $p \times q$ di antara $y_1, y_2, \dots, y_p$ dan $x_1, x_2, \dots, x_q$

S_{YY}: matriks varian sampel berukuran $p \times p$ dari $y_1, y_2, \dots, y_p$

(Rencher, 2002).

Kombinasi linear yang terbentuk oleh $y_1, y_2, \dots, y_p$ dan $x_1, x_2, \dots, x_q$ dapat dituliskan sebagai berikut:

$$U = \boldsymbol{a}^T\boldsymbol{X} = \left(a_1x_1 + a_2x_2 + \cdots + a_qx_q\right) \tag{2.10}$$

$$V = \boldsymbol{b}^T\boldsymbol{Y} = (b_1y_1 + b_2y_2 + \cdots + b_py_p) \tag{2.11}$$

dimana U dan V ialah variabel kanonik, $\boldsymbol{a}^T$ dan $\boldsymbol{b}^T$ ialah bobot kanonik atau vektor koefisien kanonik dan banyak kombinasi linear yang akan terbentuk adalah sebanyak k, yaitu nilai minimal p dan q (Dillon and Goldstein, 1984). Korelasi kanonik dapat diperoleh dari beberapa langkah berikut:

1. Varian variabel kanonik U

$$U = \boldsymbol{a}^T\boldsymbol{X} = \left(a_1x_1 + a_2x_2 + \cdots + a_qx_q\right) \tag{2.12}$$

Kombinasi linear pada percobaan ke- j dengan n sampel adalah:

$$\boldsymbol{a}^T\boldsymbol{X} = \left(a_1x_{1j} + a_2x_{2j} + \cdots + a_qx_{qj}\right);\ j = 1,2,\ldots,n \tag{2.13}$$

Maka diperoleh rata-rata sampel sebagai berikut:

$$\begin{aligned}\overline{U} &= \frac{(\boldsymbol{a}^T x_1 + \boldsymbol{a}^T x_2 + \cdots + \boldsymbol{a}^T x_n)}{n}\\ &= \boldsymbol{a}^T(x_1 + x_2 + \cdots + x_n)(\tfrac{1}{n})\\ &= \boldsymbol{a}^T\bar{x}\end{aligned} \tag{2.14}$$

Karena $(\boldsymbol{a}^T x_i - \boldsymbol{a}^T\bar{x})^2 = \boldsymbol{a}^T(x_i - \bar{x})^2 = \boldsymbol{a}^T(x_i - \bar{x})\ (x_i - \bar{x})^T\boldsymbol{a}$ maka diperoleh variansi sampel sebagai berikut:

$$\begin{aligned}Var(U) &= \frac{(\boldsymbol{a}^T x_1 - \boldsymbol{a}^T\bar{x})^2 + \cdots + (\boldsymbol{a}^T x_n - \boldsymbol{a}^T\bar{x})^2}{n-1}\\ &= \frac{(\boldsymbol{a}^T(x_1-\bar{x})\ (x_1-\bar{x})^T\boldsymbol{a} + \cdots + \boldsymbol{a}^T(x_n-\bar{x})\ (x_n-\bar{x})^T\boldsymbol{a}}{n-1}\\ &= \boldsymbol{a}^T\left[\frac{(x_1-\bar{x})\ (x_1-\bar{x})^T + \cdots + (x_n-\bar{x})\ (x_n-\bar{x})^T}{n-1}\right]\boldsymbol{a}\\ &= \boldsymbol{a}^T\boldsymbol{S_{xx}}\boldsymbol{a}\end{aligned} \tag{2.15}$$

(Dillon and Goldstein, 1984).

2. Varian Variabel Kanonik V

$$V = \boldsymbol{b}^T\boldsymbol{Y} = \left(b_1y_1 + b_2y_2 + \cdots + b_py_p\right) \tag{2.16}$$

Kombinasi linear pada percobaan ke- j dengan n sampel adalah:

$$\boldsymbol{b}^T\boldsymbol{Y} = \left(b_1y_{1j} + b_2y_{2j} + \cdots + b_py_{pj}\right);\ j = 1,2,\ldots,n \tag{2.17}$$

Maka diperoleh rata-rata sampel sebagai berikut:

$$\begin{aligned}\bar{V} &= \frac{(\boldsymbol{b}^T y_1 + \boldsymbol{b}^T y_2 + \cdots + \boldsymbol{b}^T y_n)}{n}\\ &= \boldsymbol{b}^T(y_1 + y_2 + \cdots + y_n)(\tfrac{1}{n})\\ &= \boldsymbol{b}^T\bar{y}\end{aligned} \tag{2.18}$$

Karena $(\boldsymbol{b}^T y_i - \boldsymbol{b}^T\bar{y})^2 = \boldsymbol{b}^T(y_i - \bar{y})^2 = \boldsymbol{b}^T(y_i - \bar{y})\ (y_i - \bar{y})^T\boldsymbol{b}$ maka diperolehvariansi sampel sebagai berikut:

$$\begin{aligned}Var(V) &= \frac{(\boldsymbol{b}^T y_1 - \boldsymbol{b}^T\bar{y})^2 + \cdots + (\boldsymbol{b}^T y_n - \boldsymbol{b}^T\bar{y})^2}{n-1}\\ &= \frac{\boldsymbol{b}^T(y_1-\bar{y})\ (y_1-\bar{y})^T\boldsymbol{b} + \cdots + \boldsymbol{b}^T(y_n-\bar{y})\ (y_n-\bar{y})^T\boldsymbol{b}}{n-1}\end{aligned}$$

$$= \boldsymbol{b}^T \left[\frac{(y_1-\bar{y})\,(y_1-\bar{y})^T+\cdots+(y_n-\bar{y})\,(y_n-\bar{y})^T}{n-1}\right] \boldsymbol{b}$$

$$= \boldsymbol{b}^T \boldsymbol{S}_{yy} \boldsymbol{b} \qquad (2.19)$$

(Dillon and Goldstein, 1984).

3. Kovarian Kombinasi Linear (U, V)

$$U = \boldsymbol{a}^T \boldsymbol{X} = \left(a_1 x_1 + a_2 x_2 + \cdots + a_q x_q\right) \qquad (2.20)$$

$$V = \boldsymbol{b}^T \boldsymbol{Y} = \left(b_1 y_1 + b_2 y_2 + \cdots + b_p y_p\right) \qquad (2.21)$$

$$Cov(U, V) = \frac{(\boldsymbol{a}^T x_1 - \boldsymbol{a}^T \bar{x})(\boldsymbol{b}^T y_1 - \boldsymbol{b}^T \bar{y}) + \cdots + (\boldsymbol{a}^T x_n - \boldsymbol{a}^T \bar{x})(\boldsymbol{b}^T y_n - \boldsymbol{b}^T \bar{y})}{n-1}$$

$$= \frac{\boldsymbol{a}^T (x_1-\bar{x})(y_1-\bar{y})\boldsymbol{b} + \cdots + \boldsymbol{a}^T (x_n-\bar{x})(y_n-\bar{y})\boldsymbol{b}}{n-1}$$

$$= \boldsymbol{a}^T \left[\frac{(x_1-\bar{x})(y_1-\bar{y}) + \cdots + (x_n-\bar{x})(y_n-\bar{y})}{n-1}\right] \boldsymbol{b}$$

$$= \boldsymbol{a}^T \boldsymbol{S}_{xy} \boldsymbol{b} \qquad (2.22)$$

Dari persamaan (2.15), (2.19), (2.22) dapat diperoleh korelasi kanonik:

$$\rho = Corr(U, V) = \frac{Cov(U,V)}{\sqrt{Var(U)}\sqrt{Var(V)}} = \frac{\boldsymbol{a}^T \boldsymbol{S}_{xy} \boldsymbol{b}}{\sqrt{\boldsymbol{a}^T \boldsymbol{S}_{xx} \boldsymbol{a}}\sqrt{\boldsymbol{b}^T \boldsymbol{S}_{yy} \boldsymbol{b}}} \qquad (2.23)$$

(Dillon and Goldstein, 1984).

Sehingga dapat dikatakan bahwa pasangan pertama dari variabel kanonik ialah kombinasi linear U_1, V_1 yang memiliki ragam satu dan korelasi terbesar, pasangan kedua dari variabel kanonik ialah kombinasi linear U_2, V_2 yang memiliki ragam satu dan korelasi terbesar kedua serta tidak berkorelasi dengan variabel kanonik pertama, pasangan ke- k dari variabel kanonik ialah kombinasi linear U_k, V_k yang memiliki ragam satu dan korelasi terbesar ke- k serta tidak berkorelasi dengan variabel kanonik $1,2, \ldots, k-1$. Dengan demikian dapat dituliskan (Johnson and Wichren, 2007):

1. Variabel kanonik pertama:

$$U_1 = a_1^T \boldsymbol{X} \qquad V_1 = b_1^T \boldsymbol{Y}$$

$$Var(U_1) = Var(V_1) = 1 \qquad (2.24)$$

Maksimum Corr$(U_1, V_1) = \rho_1$.

2. Variabel kanonik kedua:

$$U_2 = a_2^T \boldsymbol{X} \qquad V_2 = b_2^T \boldsymbol{Y}$$

$$Var(U_2) = Var(V_2) = 1$$

$$Cov(U_1, U_2) = Cov\left(V_{1,} V_2\right) = 0 \qquad (2.25)$$

Maksimum Corr$(U_2, V_2) = \rho_2$.

3. Variabel kanonik ke-k

$$U_k = a_k^T \boldsymbol{X} \qquad V_k = b_k^T \boldsymbol{Y}$$

$$Var(U_k) = Var(V_k) = 1$$

$$Cov(U_k, U_l) = 0 \qquad (2.26)$$

Maksimum Corr$(U_k, V_k) = \rho_k$.

Dimana, $\rho_1^2 > \rho_2^2 > \cdots > \rho_k^2$ (jika $p < q$) dan $e_1, e_2, \ldots, e_k$ adalah eigen value dan eigen vektor dari matriks $S_{xx}^{-\frac{1}{2}} S_{xy} S_{yy}^{-1} S_{yx} S_{xx}^{-\frac{1}{2}}$. Nilai $\rho_1^2 > \rho_2^2 > \cdots > \rho_k^2$ juga merupakan *p-eigen* value terbesar dari matriks $S_{yy}^{-\frac{1}{2}} S_{yx} S_{xx}^{-1} S_{xy} S_{yy}^{-\frac{1}{2}}$ dengan pasangan eigen vektornya adalah $f_1, f_2, \ldots, f_k$ (Johnson and Wichern, 2007). Nilai eigen diperoleh dari persamaan karakteristik:

$$|S_{xx}^{-1} S_{xy} S_{yy}^{-1} S_{yx} - \lambda I| = 0 \text{ dan } |S_{yy}^{-1} S_{yx} S_{xx}^{-1} S_{xy} - \lambda I| = 0 \qquad (2.27)$$

Vektor koefisien a_k dan b_k diperoleh pada fungsi kanonik $U_k = a_k^T \boldsymbol{X}$ dan $V_k = b_k^T \boldsymbol{Y}$ merupakan vektor eigen dari dua matriks yang sama:

$$\left(S_{xx}^{-1} S_{xy} S_{yy}^{-1} S_{yx} - \lambda I\right)\boldsymbol{a} = 0 \text{ dan } \left(S_{yy}^{-1} S_{yx} S_{xx}^{-1} S_{xy} - \lambda I\right)\boldsymbol{b} = 0 \qquad (2.28)$$

Sehingga, dua matriks $S_{xx}^{-1} S_{xy} S_{yy}^{-1} S_{yx}$ dan $S_{yy}^{-1} S_{yx} S_{xx}^{-1} S_{xy}$ mempunyai nilai eigen tak nol (Rencher, 2002). Untuk pasangan fungsi kanonik ke-k:

$$U_1 = a_1^T \boldsymbol{X} \qquad V_1 = b_1^T \boldsymbol{Y}$$

$$\vdots \qquad\qquad \vdots$$

$$U_k = a_k^T \boldsymbol{X} \qquad V_k = b_k^T \boldsymbol{Y} \qquad (2.29)$$

Dimana $\boldsymbol{X}$ dan $\boldsymbol{Y}$ adalah nilai-nilai dari himpunan variabel independen dan dependen. Secara umum, terdapat $k = \min(p, q)$ nilai dari korelasi kanonik ρ_k dan dengan k pasang variabel kanonik yang bersesuaian $U_k = a_k^T \boldsymbol{X}$ dan $V_k = b_k^T \boldsymbol{Y}$.

$$Corr(U_k, V_k) = \frac{Cov(U_k, V_k)}{\sqrt{Var(U_k)}\sqrt{Var(V_k)}} = \frac{a_k^T \boldsymbol{S_{xy}} \boldsymbol{b}}{\sqrt{a_k^T \boldsymbol{S_{xx}} \boldsymbol{a}}\sqrt{b_k^T \boldsymbol{S_{yy}} \boldsymbol{b}}} = \rho_k \qquad (2.30)$$

(Johnson and Wichern, 2007).

2.4 Uji Hipotesis Korelasi Kanonik

Terdapat dua hipotesis yang akan diujikan dalam analisis korelasi kanonik ialah uji hipotesis untuk mengetahui apakah secara simultan korelasi kanonik signifikan dan uji hipotesis untuk mengetahui apakah secara parsial korelasi kanonik signifikan (Rencher, 2002). Jika uji hipotesis pertama diperoleh kesimpulan bahwa korelasi kanonik signifikan maka dilanjutkan dengan uji hipotesis kedua untuk mengetahui apakah secara parsial korelasi kanonik signifikan.

a. Uji Hipotesis Secara Simultan

- Hipotesis:

  H_0: $\rho_{c1} = \rho_{c2} = \cdots = \rho_{ck} = 0$ (semua korelasi kanonik tidak signifikan)

  H_1: $\rho_{ci} > 0, i = 1,2,\ldots,k$ (paling tidak ada satu korelasi kanonik signifikan)

  Uji signifikansi yang akan digunakan ialah dengan uji *Wilk's Lambda*. *Wilk's Lambda* adalah statistik uji yang digunakan jika terdapat lebih dari satu variabel bebas. Semakin kecil nilai *Wilk's Lambda*, maka semakin tinggi kemungkinan tidak ada perbedaan rata-rata antar kedua kelompok. Sehingga terdapat hubungan yang signifikan antara kedua kelompok variabel. Nilai *Wilk's Lambda* berkisar antara 0-1. Statistik uji dirumuskan sebagai berikut.

$$\chi^2 = -\left[n - \frac{1}{2}(p + q + 3)\right] \ln \Lambda_k \qquad (2.31)$$

  dimana,

$$\Lambda_k = \prod_{i=1}^{k}(1 - \rho_i^2)$$

  dengan:

  Λ_k: koefisien *Wilk's Lambda*

  n : jumlah pengamatan

  p : jumlah variabel dependen

  q : jumlah variabel independen

- Kriteria keputusan:

  H_0 ditolak jika $\chi^2_{hit} > \chi^2_{\alpha}$ dengan derajat bebas $p \times q$ (Rencher, 2002).

b. Uji Hipotesis Secara Sebagian (Individu)

- Hipotesis:

  H_0: $\rho_{c1} = \rho_{c2} = \cdots = \rho_{ck} = 0$ (semua korelasi kanonik tidak signifikan)

  H_1: $\rho_{ci} > 0, i = 1,2,\ldots,k$ (paling tidak ada satu korelasi kanonik signifikan)

  Statistik uji Rao, pendekatan F:

$$F = \frac{1-\Lambda_k^{\frac{1}{t}}}{\Lambda_k^{\frac{1}{t}}}\left(\frac{df_2}{df_1}\right) \tag{2.32}$$

  dimana,

$$\Lambda_k = \prod_{i=1}^{k}(1-\rho_i^2)$$

$$df_1 = (p-k+1)(q-k+1)$$

$$df_2 = wt - \frac{1}{2}[(p-k+1)(q-k+1)] + 1$$

$$w = n - \frac{1}{2}(p+q+3)$$

$$t = \sqrt{\frac{(p-k+1)^2(q-k+1)^2-4}{(p-k+1)^2+(q-k+1)^2-5}}$$

  dengan:

  Λ_k: koefisien *Wilk's Lambda*

  n : jumlah pengamatan

  p : jumlah variabel terikat

  q : jumlah variabel bebas

- Kriteria keputusan:

  H_0 ditolak jika $F_{hit} > F_{\alpha,df_1,df_2}$ (Rencher, 2002).

2.5 Analisis Redundansi

Redundansi ialah sebuah ukuran yang menjelaskan besarnya keragaman yang dijelaskan berdasarkan korelasi antara variabel dependen dan variabel independen dengan variabel kanonik (Rencher, 2002).

Besarnya keragaman untuk himpunan y yang diterangkan oleh $v_1, v_2, \ldots, v_k$:

$$Rd(y|v) = \frac{\sum_{j=1}^{P} R_{yi}^2|x}{p} \tag{2.33}$$

dengan:

$R_{yi}^2|x$: nilai korelasi antara variabel terikat (y) dengan variabel kanonik v yang dibentuk dari variabel bebas (x)

p : banyaknya variabel terikat

Indeks redundansi y yang diterangkan oleh $v_1, v_2, \ldots, v_k$:

$$RI(y|v) = Rd(y|v)\rho_{ck}^2 \quad (2.34)$$

dengan:

ρ_{ck}^2: kuadrat dari korelasi kanonik

Besarnya keragaman untuk himpunan x yang diterangkan oleh $u_1, u_2, \ldots, u_k$:

$$Rd(x|u) = \frac{\sum_{j=1}^{q} R_{xi}^2|y}{q} \quad (2.35)$$

dengan:

$R_{xi}^2|y$: nilai korelasi antara variabel bebas (x) dengan variabel kanonik u yang dibentuk dari variabel terikat (y)

p : banyaknya variabel independen

Indeks redundansi x yang diterangkan oleh $u_1, u_2, \ldots, u_k$:

$$RI(x|u) = Rd(x|u)\rho_{ck}^2 \quad (2.36)$$

dengan:

ρ_{ck}^2: kuadrat dari korelasi kanonik

(Rencher, 2002).

Untuk menentukan fungsi kanonik yang dianggap cukup dalam menjelaskan struktur hubungan Y dan X dilihat dari koefisien R^2 (R-Square). Koefisien R^2 paling tinggi dipilih untuk menetukan fungsi kanonik mana yang digunakan. Nilai ini diperoleh dengan mengkuadratkan korelasi kanonik.

$$R_k^2 = \rho_k^2 \quad (2.37)$$

Indeks redundansi dihitung untuk kedua variabel terikat dan bebas (Hair, *et al.*, 2009).

2.6 Interpretasi Fungsi Kanonik

Interpretasi yang dilakukan dalam analisis korelasi kanonik ialah terhadap bobot kanonik (*Canonical Weight*), muatan kanonik (*cannonical loading*), serta muatan silang kanonik (*cross cannonical loading*) (Mattjik dan Sumertajaya, 2011).

1. Bobot Kanonik

   Bobot kanonik di interpretasikan sebagai besarnya kontribusi variabel asal terhadap variabel kanonik (Hair, *et all.*, 2009). Semakin besar bobot kanonik maka semakin besar kontribusi variabel yang bersangkutan terhadap variabel kanonik dan begitupun sebaliknya. Begitu juga dengan variabel yang memiliki bobot nilai dengan tanda berlawanan menggambarkan hubungan kebalikan dengan variabel lainnya dan variabel dengan tanda yang sama menunjukkan hubungan langsung (Hair, *et all.*, 2009).

2. Muatan Kanonik

   Muatan kanonik saat ini banyak dipakai untuk interpretasi. Muatan kanonik dapat disebut korelasi struktur kanonik. Muatan kanonik dihitung dari korelasi antara variabel asal dengan masing-masing variabel kanoniknya. Semakin besar nilai muatan kanonik menunjukkan semakin dekat hubungan fungsi kanonik yang bersangkutan dengan variabel asal (Mattjik dan Sumertajaya, 2011).

a. Muatan kanonik variabel bebas:

$$R_{xu} = R_{xx}\boldsymbol{a} \tag{2.38}$$

dengan:

R_{xu}: muatan kanonik himpunan variabel bebas (X)

R_{xx}: matriks korelasi antar variabel bebas (X)

$\boldsymbol{a}$: vektor koefisien kanonik variabel u

b. Muatan kanonik variabel terikat:

$$R_{yv} = R_{yy}\boldsymbol{b} \tag{2.39}$$

dengan:

R_{yv}: muatan kanonik himpunan variabel terikat (Y)

R_{yy}: matriks korelasi antar variabel terikat (Y)

$\boldsymbol{b}$: vektor koefisien kanonik variabel v

3. Muatan Silang Kanonik

Muatan silang kanonik dihitung dari korelasi antara variabel asal terhadap variabel kanonik lainnya. Semakin besar nilai muatan silang kanonik mencerminkan semakin dekat hubungan fungsi kanonik yang bersangkutan dengan variabel asal (Mattjik dan Sumertajaya, 2011).

a. Muatan silang kanonik variabel bebas:

$$R_{xv} = R_{xu}\rho_k \tag{2.40}$$

dengan:

R_{xv}: muatan silang kanonik himpunan variabel bebas (X)

R_{xu}: muatan kanonik himpunan variabel bebas (X)

ρ_k : nilai korelasi kanonik dari variabel kanonik ke-k

b. Muatan silang kanonik variabel terikat:

$$R_{yu} = R_{yv}\rho_k \tag{2.41}$$

dengan:

R_{yu}: muatan silang kanonik himpunan variabel terikat (Y)

R_{yv}: muatan kanonik himpunan variabel terikat (Y)

ρ_k : nilai korelasi kanonik dari variabel kanonik ke-k

2.7 Pendekatan *Robust* untuk Analisis Korelasi Kanonik

Menurut Huber (1981), metode pendugaaan *robust* mendapat perhatian yang cukup menarik, karena mempunyai keterkaitan besar dengan masalah dan studi pendugaan lokasi yang sensitif terhadap pencilan. Vektor rata-rata sampel dan matriks varian kovarian sampel merupakan landasan dasar dalam analisis multivariat. Keduanya akan mendapatkan hasil yang optimal jika data berdistribusi normal.

Distribusi normal mempunyai peran yang penting dalam analisis multivariat. Vektor rata-rata sampel dan matriks varian kovarian sampel akan menjadi kurang efisien ketika terdapat pencilan pada data yang akan dianalisis. Jarak mahalanobis belum mampu untuk mengatasi pencilan dengan jumlah lebih dari satu. Selanjutnya terdapat beberapa pengembangan teori untuk mengatasi pencilan pada kasus multivariat, salah satunya menggunakan pendekatan *robust* multivariat. Salah satu pendekatan *robust* pada analisis korelasi kanonik menggunakan matriks varian kovarian dengan penduga *Minimum Covariance Determinant* (MCD). Penduga *robust* MCD ialah rata-rata dan varian kovarian dari sebagian pengamatan yang meminimumkan determinan matriks varian kovarian.

2.7.1 Deteksi Pencilan Multivariat

Pencilan ialah pengamatan yang tidak mengikuti sebagian besar pola dan terletak jauh dari pusat data (Barnett and Lewis, 1978). Jarak mahalanobis menyatakan bahwa pengamatan ke-i didefinisikan sebagai pencilan jika jaraknya lebih besar dari nilai *chi square* tabel pada sejumlah pengamatan. Perhitungan jarak mahalanobis sebagai berikut:

$$d_{MD}^2 = (x_i - \overline{\boldsymbol{x}})^T \boldsymbol{S}^{-\mathbf{1}} (x_i - \overline{\boldsymbol{x}}) > \chi^2_{P,(1-\alpha)}\ , i = 1,2,\dots,n \quad (2.42)$$

dengan:

$\overline{\boldsymbol{x}}$: vektor rata-rata
$\boldsymbol{S}$: matriks varian kovarian
p : banyaknya variabel pengamatan
$x_1, x_2, \dots, x_n$: observasi sampel

Jarak mahalanobis dapat digunakan untuk mendeteksi adanya pencilan tunggal dalam data. Suatu hal yang sulit untuk melakukan pendeteksian adanya pencilan pada p variabel data, dimana $p > 2$. Jika terdapat pencilan tunggal kita dapat menggunakan jarak mahalanobis, namun pendekatan ini kurang efektif untuk pengamatan dengan pencilan lebih dari satu. Identifikasi pencilan menjadi kurang optimal karena adanya pengaruh *masking* (penyamaran) dan *swamping* (pelimpahan). Penyamaran terjadi pada saat pengamatan pencilan tidak terdeteksi karena adanya pengamatan pencilan lain yang berdekatan, sedangkan Pelimpahan terjadi pada saat pengamatan baik teridentifikasi sebagai pengamatan pencilan.

Oleh karena itu, sebaiknya menggunakan perhitungan jarak dengan penduga *robust* bagi sebaran multivariat. Jarak *robust* ialah suatu pendekatan untuk mengidentifikasi pencilan pada data multivariat, dengan menggunakan penaksir dari vektor rata-rata ($\bar{x}$) dan matriks varian kovarian ($\boldsymbol{S}$). Sehingga, metode ini mampu meminimumkan pengaruh dari adanya efek *masking* (penyamaran) dan *swamping* (pelimpahan) dalam pendeteksian pencilan (Rencher, 2002). Salah satu penduga *robust* yang mempunyai kemampuan mengukur jarak dan mendeteksi titik *leverage* (pencilan yang disebabkan oleh variabel independen) adalah *Minimum Covariance Determinant* (MCD). Sebuah pengamatan x_i diidentifikasi sebagai pencilan jika jarak mahalanobis *robust* sebagai berikut:

$$d_{RD}^2 = (x_i - \bar{\boldsymbol{x}}_{\boldsymbol{MCD}})^T \boldsymbol{S}_{\boldsymbol{MCD}}^{-\mathbf{1}} (x_i - \bar{\boldsymbol{x}}_{\boldsymbol{MCD}}) > \chi^2_{P,(1-\alpha)}, i = 1,2, \dots, n \qquad (2.43)$$

Dimana $\bar{\boldsymbol{x}}_{\boldsymbol{MCD}}$ dan $\boldsymbol{S}_{\boldsymbol{MCD}}$ menyatakan vektor rata-rata dan matriks varian kovarian dari sebagian data X yang mempunyai determinan matriks varian kovarian terkecil (Hubert, *et all.*, 2008).

2.7.2 *Minimum Covariance Determinant* (MCD)

Metode *Minimum Covariance Determinant* (MCD) adalah penduga *robust* untuk rata-rata dan matriks varian kovarian dengan mencari himpunan data yang menghasilkan determinan terkecil yang digunakan untuk mendeteksi pencilan (Rousseeuw and Driessen, 1999). Metode ini bertujuan untuk mendapatkan suatu sub sampel berukuran h dari keseluruhan pengamatan n yang matriks varian kovariannya memiliki determinan terkecil diantara semua kemungkinan kombinasi data. Misalkan himpunan bagian tersebut x_h, maka terdapat C_h^n kombinasi himpunan yang harus dicari untuk mendapatkan penduga MCD dengan,

$$h = \frac{(n+p+1)}{2}, h \leq n \qquad (2.44)$$

dimana p menyatakan banyaknya variabel (Rousseeuw and Driessen, 1999).

Misalkan $\boldsymbol{X} = (x_1, x_2, \dots, x_n)^T$ merupakan himpunan data sejumlah n pengamatan terdiri dari p- variabel dimana $n \geq p + 1$. Penduga MCD merupakan pasangan sub sampel $\boldsymbol{T}$ dan $\boldsymbol{S}$ yang termasuk matriks definit positif simetri berdimensi $p \times p$ dari suatu sub sampel berukuran h pengamatan dimana $h = \frac{(n+p+1)}{2}, h \leq n$ dengan

$$\boldsymbol{T} = \frac{1}{h}\sum_{i=1}^{h} x_i \text{ , } i = 1,2, \dots, n \tag{2.45}$$

$$\boldsymbol{S} = \frac{1}{h}\sum_{i=1}^{h}(x_i - \boldsymbol{T_i})^T(\, x_i - \boldsymbol{T}_i) \text{ , } \; i = 1,2, \dots, n \tag{2.46}$$

dimana,

$$h = \frac{(n+p+1)}{2}$$

dengan:

$\boldsymbol{T}$: vektor rata-rata

$\boldsymbol{S}$: matriks varian kovarian

$x_1, x_2, \dots, x_n$: observasi sampel

(Rousseeuw and Driessen, 1999).

Pada populasi dengan jumlah pengamatan yang kecil, penduga MCD dapat dengan cepat dihitung. Tetapi jika jumlah pengamatan besar, maka akan banyak sekali kombinasi sub-sampel dari H yang harus dicari dan perhitungannya akan cukup memakan waktu. Dalam mengatasi keterbatasan ini, maka digunakan suatu algoritma baru untuk metode MCD yang dinamakan dengan metode *fast*-MCD yaitu teorema *C-Step* (Dayanti, dkk, 2016).

Teorema *C-Step*

Menurut Rousseeuw and Driessen (1999), misalkan $\boldsymbol{X} = (x_1, x_2, \dots, x_n)^T$ ialah himpunan sejumlah n pengamatan terdiri dari p- variabel. Misalkan $H_1 \subset \{1,2, \dots, n\}$ dengan sejumlah elemen $|H_1| = h$, tetapkan $T_1 = \frac{1}{h}\sum_{i=1}^{h} x_i$ dan $S_1 = \frac{1}{h}\sum_{i=1}^{h}(x_i - \boldsymbol{T}_1)^T(\, x_i - \boldsymbol{T}_1)$. Jika det $(S_1) \neq 0$ maka jarak relatifnya:

$$d_1(i) = \sqrt{(x_i - T_1)^T S_1^{-1}(x_i - T_1)} \text{ , } i = 1,2, \dots, n \tag{2.47}$$

Selanjutnya, ambil H_2 sedemikian sehingga

$\{d_1(i); i \in H_2\} \coloneqq \{(d_1)_{1:n}, (d_1)_{2:n}, \dots, (d_1)_{h:n}\}$ dimana

$(d_1)_{1:n} \leq (d_1)_{2:n} \leq \cdots \leq (d_1)_{h:n}$ adalah orde jarak T_2 dan S_2 dihitung berdasarkan H_2. Selanjutnya, $\det(S_2) \leq \det(S_1)$ dan akan sama jika dan hanya jika $T_1 = T_2$ dan $S_1 = S_2$ (Rousseeuw and Driessen, 1999).

Bukti:

Asumsikan bahwa $\det(S_2) > 0$, selanjutnya jarak relatif $d_2(i) = d_{(T_2,\, S_2)}(i)$ untuk semua $i = 1,2,\dots,n$ dengan menggunakan $|H_2| = h$ didefinisikan dari (T_2, S_2) maka diperoleh:

$$\begin{aligned}\frac{1}{hp}\sum_{i\epsilon H_2} d_2^2(i) &= \frac{1}{hp} tr \sum_{i\epsilon H_2} (x_i - T_2)^T S_2^{-1}(x_i - T_2)\\ &= \frac{1}{hp} tr \sum_{i\epsilon H_2} S_2^{-1}(x_i - T_2)^T (x_i - T_2)\\ &= \frac{1}{p} tr S_2^{-1} S_2\\ &= \frac{1}{p} tr(I) = 1 \qquad (2.48)\end{aligned}$$

Selanjutnya,

$$\lambda \coloneqq \frac{1}{hp}\sum_{i\epsilon H_2} d_1^2(i) = \frac{1}{hp}\sum_{i=1}^{h} (d_1^2)_{i:n} \leq \frac{1}{hp}\sum_{j\epsilon H_1} d_1^2(j) = 1 \qquad (2.49)$$

Dengan $\lambda = 0$, sebab jika tidak det $(S_2) = 0$. Kombinasikan persamaan (2.48) dan pertidaksamaan (2.49) dihasilkan:

$$\begin{aligned}\frac{1}{hp}\sum_{i\epsilon H_2} d^2_{(T_1,\lambda S_1)}(i) &= \frac{1}{hp}\sum_{i\epsilon H_2} (x_i - T_1)^T \frac{1}{\lambda} S_1^{-1}(x_i - T_1)\\ &= \frac{1}{\lambda hp}\sum_{i\epsilon H_2} d_1^2(i)\\ &= \frac{\lambda}{\lambda} = 1 \qquad (2.50)\end{aligned}$$

Bukti bahwa (T_2, S_2) merupakan ukuran minimum pada $\det(\boldsymbol{S})$ diantara semua $(\boldsymbol{T}, \boldsymbol{S})$ untuk $\frac{1}{hp}\sum_{i\epsilon H_2} d^2_{(T,S)}(i) = 1$ ini mengakibatkan bahwa $\det(S_2) \leq \det(\lambda S_1)$. Pada pertidaksamaan (2.49) telah disebutkan bahwa $\det(\lambda S_1) \leq \det(S_1)$, sehingga

$$\det(S_2) \leq \det(\lambda S_1) \leq \det(S_1). \qquad (2.51)$$

Dengan catatan bahwa $\det(S_2) = \det(S_1)$ jika dan hanya jika diantara pertidaksamaan (2.51) adalah sama. Sehingga, dapat diambil kesimpulan bahwa $\det(S_2) = \det(\lambda S_1)$ jika dan hanya jika $(T_2, S_1) = (T_1, \lambda S_1)$. Kedua, $\det(\lambda S_1) = \det(S_1)$ jika dan hanya jika $\lambda = 1$ hal itu berarti $S_1 = \lambda S_1$. Kombinasi keduanya menjadi $(T_2, S_2) = (T_1, S_1)$.

Algoritma *fast*-MCD adalah sebagai berikut (Rousseeuw and Driessen, 1999):

1. Ambil himpunan $\boldsymbol{X}$ secara acak. Misalkan himpunan bagian tersebut sebagai H_1 dimana $h = \frac{(n+p+1)}{2}$.
2. Hitung vektor rata-rata T_1 dan matriks varian kovarian S_1 dari H_1 dengan $T_1 = \frac{1}{h}\sum_{i=1}^{h} x_i$ dan $S_1 = \frac{1}{h}\sum_{i=1}^{h}(x_i - \boldsymbol{T_1})^T(x_i - \boldsymbol{T_1})$.
3. Hitung determinan dari matriks varian kovarian S_1.
4. Hitung jarak relatif dari setiap pengamatan terhadap rata-rata dan matriks varian kovarian dengan persamaan:

$$d_1(i) = \sqrt{(x_i - T_1)^T S_1^{-1}(x_i - T_1)} \tag{2.52}$$

5. Urutkan jarak relatif berdasarkan jarak mahalanobis dari yang terkecil hingga yang terbesar.
6. Bentuk himpunan bagian baru dengan H_2 sedemikian sehingga $\{d_1(i); i \in H_2\} \coloneqq \{(d_1)_{1:n}, (d_1)_{2:n}, \dots, (d_1)_{h:n}\}$ dimana $(d_1)_{1:n} \le (d_1)_{2:n} \le \cdots \le (d_1)_{h:n}$.
7. Hitung vektor rata-rata T_2, matriks varian kovarian S_2, dan $d_2(i)$ dari H_2.
8. Mengulangi langkah 1 sampai dengan 6 hingga didapatkan bahwa $\det(S_2) \le \det(S_1)$.

2.8 Pertumbuhan Ekonomi

Pertumbuhan ekonomi ialah suatu penilaian yang berguna untuk melihat pengaruh yang disebabkan dari kebijakan pembangunan khususnya dalam bidang ekonomi (Avatara, 2013). Pertumbuhan ekonomi menjadi masalah ekonomi jangka panjang, serta fenomena penting yang dialami dunia belakangan ini. Pertumbuhan ekonomi Indonesia setiap tahunnya mengalami pertumbuhan yang naik turun. Di sisi lain, pertumbuhan ekonomi juga ialah indikator dari keberhasilan tercapainya pembangunan ekonomi yang baik yang ditandai dengan peningkatan kesejahteraan masyarakat.

Pembangunan ekonomi ialah serangkaian usaha dan kebijakan yang dilakukan oleh pemerintah suatu negara atau daerah untuk meningkatkan kesejahteraan masyarakat, memperluas lapangan pekerjaan, serta pemerataan tingkat pendapatan. Ekonomi suatu negara dikatakan bertumbuh jika kegiatan ekonomi masyarakat berdampak langsung terhadap kenaikan produksi barang dan jasanya. Pertumbuhan ekonomi dapat dilihat dari persentase penduduk miskin, tingkat pengangguran terbuka, indeks pembangunan manusia, dan persentase produk dosmetik regional bruto atas dasar harga berlaku.

2.9 Kesejahteraan Masyarakat

Penunjang kesejahteraan masyarakat mencakup delapan sektor berbeda ialah kependudukan, kesehatan dan gizi, pendidikan, ketenagakerjaan, taraf dan pola konsumsi, perumahan dan lingkungan, kemiskinan, dan sosial lainnya (BPS, 2021). Indonesia selalu berusaha demi terwujudnya ekonomi yang baik dan dapat meningkatkan kesejahteraan masyarakat. Pada dasarnya pencapaian kesejahteraan masyarakat dilakukan dengan berbagai perubahan dalam pembangunan masyarakat yang berguna untuk perbaikan kondisi ekonomi, sosial, dan budaya.

Kesejahteraan masyarakat ialah kondisi terpenuhnya kebutuhan material, spiritual, serta sosial warga negara agar dapat hidup layak dan mampu mengembangkan diri sehingga dapat melaksanakan fungsi sosialnya. Masyarakat dapat disebut sejahtera apabila masyarakat tersebut telah dapat memenuhi kebutuhan hidupnya secara mandiri. Kesejahteraan masyarakat dapat dilihat dari persentase perempuan melahirkan dibantu tenaga medis, persentase penduduk yang memiliki jaminan kesehatan, persentase rumah tangga dengan sumber penerangan listrik, persentase status kepemilikan rumah sendiri, persentase rumah tangga dengan sumber air minum yang layak, dan persentase penduduk.

III METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilaksanakan pada semester genap tahun akademik 2021/2022 dan bertempat di Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung.

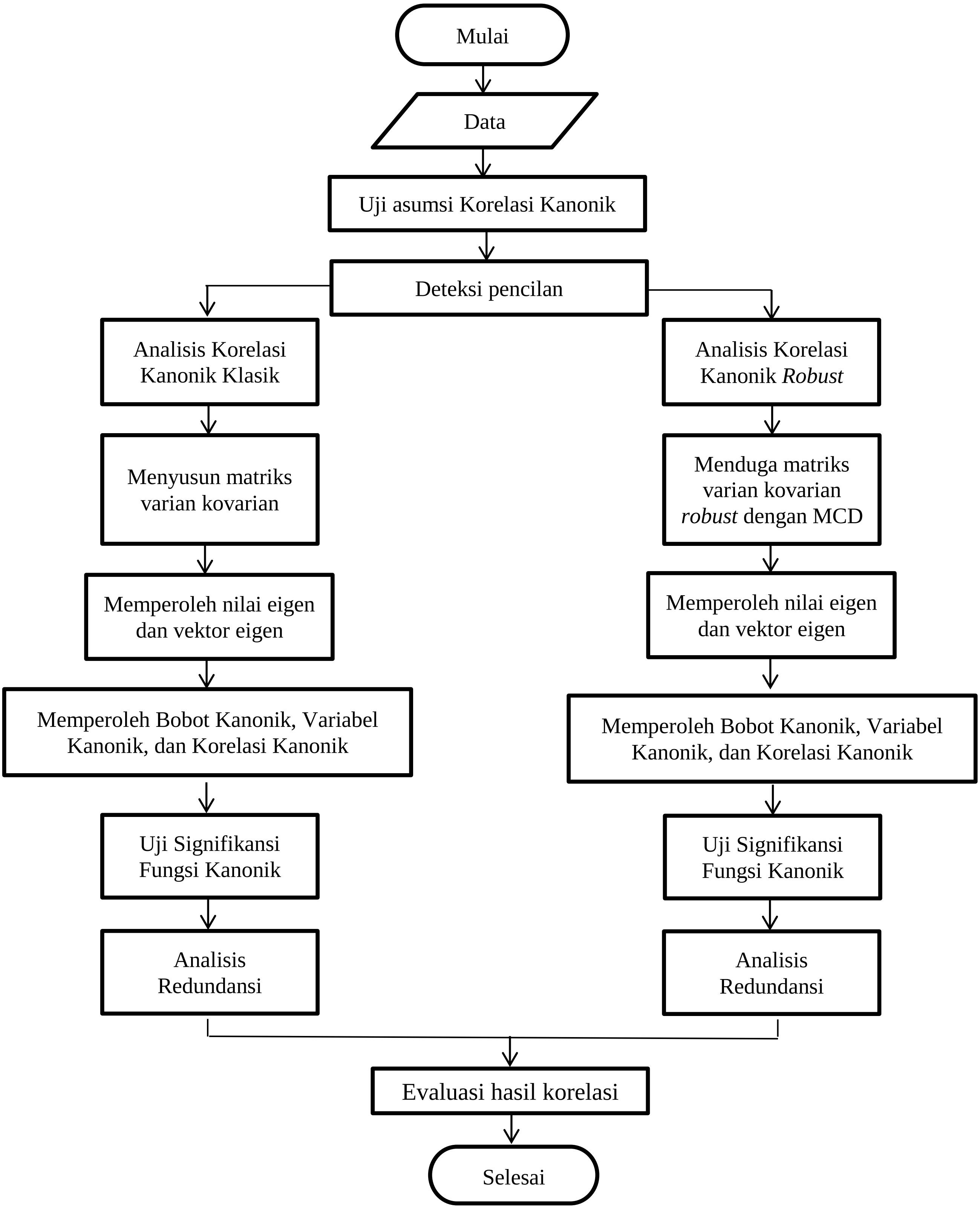
3.2 Data Penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) Indonesia yang merupakan data kesejahteraan rakyat dan pertumbuhan ekonomi setiap provinsi di Indonesia pada tahun 2020 dengan jumlah data sebanyak 34 data. Variabel terikat pada penelitian ini adalah tingkat pertumbuhan ekonomi, meliputi persentase penduduk miskin (y_1), tingkat pengangguran terbuka (y_2), indeks pembangunan manusia (y_3), dan persentase produk dosmetik regional bruto atas dasar harga berlaku (y_4). Sedangkan variabel bebas pada penelitian ini adalah indeks kesejahteraan masyarakat, meliputi persentase perempuan melahirkan dibantu tenaga medis (x_1), persentase penduduk yang memiliki jaminan kesehatan (x_2), persentase rumah tangga dengan sumber penerangan listrik (x_3), persentase status kepemilikan rumah sendiri (x_4), persentase rumah tangga dengan sumber air minum yang layak (x_5), dan persentase penduduk (x_6).

3.3 Metode Penelitian

Penelitian ini dibantu dengan menggunakan *software* R versi 4.0.3. Adapun langkah-langkah yang dilakukan pada penelitian ini adalah sebagai berikut:

1. Menginput data penelitian.
2. Melakukan uji asumsi korelasi kanonik.
3. Pendeteksian pencilan pada keseluruhan data pengamatan menggunakan jarak mahalanobis dan juga menggunakan jarak *robust*.
4. Menyusun matriks varian kovarian untuk metode korelasi kanonik, dimana matriks varian kovarian $\boldsymbol{S}$ digunakan apabila data yang diolah memiliki satuan yang sama.
5. Menduga matriks varian kovarian dengan metode *Minimun Covariance Determinant* (MCD) untuk mendapatkan matriks pada korelasi kanonik *robust*.
6. Memperoleh nilai eigen dan vektor eigen berdasarkan matriks varian kovarian.
7. Memperoleh variabel kanonik dan bobot kanonik dari vektor eigen yang diperoleh pada langkah 6.
8. Menentukan nilai penduga koefisien korelasi kanonik dan fungsi kanonik, dimana fungsi kanonik terbentuk mengikuti $\min(p, q)$ dalam setiap variabel.
9. Menguji Signifikansi korelasi kanonik secara keseluruhan dan secara sebagian. Jika fungsi kanonik tidak signifikan, maka hubungan antar variabel tidak akan diinterpretasikan.
10. Melakukan analisis redundansi untuk menghitung proporsi keragaman yang dapat dijelaskan oleh variabel kanonik.
11. Evaluasi hasil korelasi dengan metode korelasi kanonik klasik dan korelasi kanonik *robust*.



Gambar 1. Diagram Alir Penelitian

V. KESIMPULAN

5.1 Kesimpulan

Berdasarkan hasil analisis data yang telah diperoleh pada bab sebelumnya, maka kesimpulan pada penelitian ini adalah sebagai berikut:

1. Analisis korelasi kanonik *robust* dengan metode *Minimum Covariance Determinant* (MCD) dapat mendeteksi adanya pencilan pada data indeks kesejahteraan masyarakat terhadap tingkat pertumbuhan ekonomi di Indonesia. Menggunakan jarak *robust* MCD terdapat sebanyak 12 dari total 34 pengamatan yang tergolong pencilan, yaitu data ke-4, 11, 12, 13, 14, 15, 16, 17, 19, 23, 33, dan 34.
2. Berdasarkan penelitian yang dilakukan pada data indeks kesejahteraan masyarakat terhadap tingkat pertumbuhan ekonomi di Indonesia diperoleh koefisien korelasi kanonik klasik dengan nilai korelasi $\rho_1 = 0.9397$, $\rho_2 = 0.8143$, $\rho_3 = 0.4660$, dan $\rho_4 = 0.2670$. Sedangkan koefisien korelasi kanonik *robust* diperoleh dengan nilai korelasi $\rho_1 = 0.9879$, $\rho_2 = 0.8602$, $\rho_3 = 0.7012$, dan $\rho_4 = 0.5768$, dimana nilai korelasi kanonik *robust* pada penelitian ini lebih besar dibandingkan dengan nilai korelasi kanonik klasik.

   Pada nilai proporsi keragaman, hasil analisis menunjukkan bahwa variabel independen oleh variabel kanonik dependen dapat menjelaskan jumlah keragaman kanonik klasik sebesar 40.46%. Sedangkan, jumlah keragaman kanonik *robust* sebesar 43.33%. Sebaliknya, pada nilai proporsi keragaman, hasil analisis menunjukkan bahwa variabel dependen oleh variabel kanonik independen dapat menjelaskan jumlah keragaman kanonik klasik sebesar 64.05%. Sedangkan, jumlah keragaman kanonik *robust* sebesar 76.55%. Hal ini berarti bahwa proporsi keragaman kanonik *robust* mampu meningkatkan nilai proporsi keragaman dibandingkan dengan proporsi keragaman kanonik klasik.

Interpretasi korelasi kanonik dilihat dari muatan kanonik, dimana hanya fungsi kanonik pertama yang diinterpretasikan lebih lanjut karena memiliki nilai muatan kanonik terbesar. Hubungan yang terjadi adalah peningkatan persentase penduduk mengindikasi adanya peningkatan persentase PDRB atas dasar harga berlaku. Dengan demikian, dapat diartikan semakin banyak jumlah penduduk maka akan semakin banyak Sumber Daya Manusia (SDM) yang akan dibutuhkan dalam dunia kerja, hal ini akan berpengaruh terhadap peningkatan persentase PDRB atas dasar harga berlaku. Sedangkan, penurunan persentase status kepemilikan rumah sendiri mengindikasi adanya peningkatan TPT. Dengan demikian, dapat diartikan semakin sedikitnya penduduk yang belum memiliki status kepemilikan rumah sendiri maka akan semakin banyak penduduk yang pengangguran, hal ini akan berpengaruh terhadap peningkatan TPT.

DAFTAR PUSTAKA

Anderson, T.W. 2003. *An Introduction to Multivariate Statistical Analysis*. 3th Edition. John Wiley & Sons, New Jersey.

Asbah, M.F., Sudarno., & Safitri, D. 2013. Penentuan Koefisien Korelasi Kanonik dan Interpretasi Fungsi Kanonik Multivariat. *Jurnal Gaussian*. **2**(2): 119-128.

Avatara, A. 2013. Faktor-faktor yang Mempengaruhi Pertumbuhan Ekonomi di Indonesia Tahun 1992-2011 (Kajian dari Sisi Fiskal). **2**(2): 277-293.

Badan Pusat Statistik. 2021. *Statistik Indonesia*. Badan Pusat Statistik, Jakarta.

Barnett, V & Lewis, T. 1978. *Outliers in Statistical Data.* John Wiley & Sons, New York.

Barrera, M.S. & Yohai, V.J. 2006. A Fast Algorithm for S-Regression Estimates. *Journal of Computational and Graphical Statistics*. **15**(2): 414-427.

Bilodeau, M. & Brenner, D. 1999. *Theory of Multivariate Statistics*. Springer, New York.

Dayanti, N.P., Suciptawati, N.L., & Susilawati, M. 2016. Penerapan Bootstrap dalam Metode Minimum Covariance Determinant (MCD) dan Least Median Square (LMS) pada Analisis Regresi Linear Berganda. *E-Jurnal Matematika*. **5**(1): 22-26.

Dillon, W.R. & Goldstein, M. 1984. *Multivariate Analysis: Methods and Applications*. John Wiley & Sons, Canada.

Hair, J.F., Black, W.C., Babin, B.J., & Anderson, R.E. 2009. *Multivariate Data Analysis*. 7th Edition. Prentice Hall, New Jersey.

Huber, P.J. 1981. *Robust Statistics*. John Wiley & Sons, New York.

Hubert, M. & Debruyne, M. 2010. Minimum Covariance Determinant. WIREs Computational Statistics. **2**(1): 36-43.

Hubert, M., Rousseeuw, P.J., & Aelst, S.V. 2008. High-Breakdown Robust Multivariate Methods. *Statistical Science*. **23**(1): 92-119.

Johnson, R.A. & Wichern, D.W. (2007). *Applied Multivariate Statistical Analysis*. 6th Edition. Prentice Hall, New Jersey.

Lopuhaa, H.P. & Rousseeuw, P.J. 1991. Breakdown Points of Affine Equivariant Estimators of Multivariate Location and Covariance Matrices. *The Annals of Statistics*. **19**(1): 229-248.

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

Nugroho, S. 2008. *Statistika Multivariat Terapan*. 1th Edition. UNIB Press, Bengkulu.

Rencher, A.C. 2002. *Methods of Multivariate Analysis*. 2th Edition. John Wiley & Sons, New York.

Rousseeuw, P.J. & Driessen, K.V. 1999. A Fast Algorithm for the Minimum Covariance Determinant Estimator. *Technometrics Journal*. **41**(3): 212-223.

Wustqa, D.U., Listyani, E., Subekti, R., Kusumawati, R., Susanti, M., & Kismiantini. 2018. Analisis Data Multivariat Dengan Program R. *Jurnal Pengabdian Masyarakat MIPA dan Pendidikan MIPA*. **2**(2): 83-86.