

BAB I

PENDAHULUAN

1.1. Latar Belakang

Di lapangan, tidak sedikit dijumpai suatu kasus di mana data memiliki lebih dari satu variabel terikat dan lebih dari satu variabel bebas. Beberapa kasus dalam penelitian yang melibatkan hubungan antara dua kelompok variabel, misalnya hubungan antara sekelompok variabel kepribadian dan sekelompok variabel kemampuan, hubungan antara indeks harga dan indeks produksi. Di samping hubungan fungsional yang dinyatakan dengan persamaan regresi, ada juga yang perlu dipersoalkan yaitu ukuran kuat lemahnya antara dua kelompok variabel. Ukuran dari kuat lemahnya dua kelompok variabel dapat diukur dengan suatu teknik, yaitu dengan Analisis Korelasi Kanonik.

Analisis korelasi kanonik adalah salah satu teknik analisis statistik, yang digunakan untuk melihat hubungan antara himpunan variabel bebas (X_1, X_2, \dots, X_p) dengan himpunan variabel terikat (Y_1, Y_2, \dots, Y_q) . Johnson dan Winchern (2007: 539) mengemukakan bahwa analisis korelasi kanonik berfokus pada korelasi antara kombinasi linier dari himpunan variabel terikat dengan kombinasi linier dari himpunan variabel bebas. Ide utama dari analisis ini adalah mencari pasangan dari kombinasi linier ini yang memiliki korelasi terbesar. Analisis korelasi kanonik bertujuan untuk mengukur tingkat keeratan hubungan antara himpunan variabel bebas dengan himpunan variabel terikat.

Interpretasi yang dapat dilakukan dalam analisis korelasi kanonik yaitu terhadap koefisien kanonik (pembobot kanonik / *canonical weight*), *canonical loadings* dan *canonical cross loadings*. Bobot kanonik merupakan koefisien kanonik yang telah dibakukan, dapat diinterpretasikan sebagai besarnya kontribusi variabel asal terhadap variabel kanonik. Semakin besar nilai koefisien ini maka semakin besar kontribusi variabel yang bersangkutan terhadap variabel kanonik. *Canonical loadings* (R_{UY}, R_{VX}) dapat dihitung dari korelasi antar variabel asal dengan masing-masing variabel kanoniknya. Semakin besar nilai *loadings* mencerminkan semakin dekat hubungan pasangan variabel kanonik yang bersangkutan dengan variabel asal. *Canonical Cross loadings* (R_{UX}, ρ_{VY}) dapat dihitung dari korelasi antara variabel asal dengan bukan variabel kanoniknya. Semakin besar nilai *loadings* mencerminkan semakin dekat hubungan pasangan variabel kanonik yang bersangkutan dengan variabel asal.

Adapun asumsi-asumsi dalam analisis korelasi kanonik yang harus dipenuhi adalah sebagai berikut :

- a. Variabel bebas dan variabel terikat terdiri atas lebih dari satu variabel dan berskala interval
- b. Tidak terdapat multikolinieritas yang sangat tinggi atau sempurna antar variabel bebasnya dan antar variabel terikatnya
- c. Berdistribusi Normal Multivariat, baik variabel terikat maupun variabel bebasnya
- d. Linieritas, yaitu semua variabel bebas mempengaruhi secara searah terhadap semua variabel terikat

Selain asumsi-asumsi di atas, perlu diperhatikan juga dalam data tersebut memiliki pencilan atau tidak. Hal ini dikarenakan secara umum analisis korelasi kanonik ditaksir berdasarkan penaksir dari matriks kovarians, yaitu *Maximum Likelihood Estimator* (MLE), perhitungannya didasarkan pada matriks variansi-kovariansi sampel biasa, yaitu $S_{ik} = \frac{1}{(n-1)} \sum_{j=1}^n (W_{ji} - \bar{W}_i)(W_{jk} - \bar{W}_k)$, $i = k = 1, 2, \dots, d$ sebagai elemen matriks S dengan vektor acak X dan Y dapat dianggap bersama-sama sebagai variabel acak W , dimensi $d = p + q$. Penggunaan matriks variansi-kovariansi ini sangat sensitif terhadap adanya pencilan. Menurut Yohai (2006: xvi) jika data mengandung pencilan maka akan mengakibatkan kenormalan dari distribusi data menjadi terganggu karena penyebaran data yang sangat jauh, akibatnya matriks variansi-kovariansi akan kehilangan efisiensinya dan mengakibatkan sifat penaksir menjadi bias dan tak konsisten. Oleh karena itu penggunaan metode MLE untuk kasus tersebut menjadi kurang dipercaya.

Menurut Rencher (2002: 100) ada dua pendekatan utama untuk menangani data dengan pencilan. Pertama adalah identifikasi, biasanya melibatkan penghapusan pencilan, tetapi pencilan yang dihapus tadi bisa saja dapat memberikan informasi penting tentang model atau data. Metode kedua adalah akomodasi, di mana metode analisis atau model tersebut dimodifikasi. Salah satu pendekatan untuk identifikasi pencilan multivariat atau akomodasi adalah dengan menggunakan metode *robust*. Metode tersebut meminimalkan pengaruh pencilan dalam estimasi atau kesesuaian model dan kokoh terhadap pencilan. Sehingga, tujuan dari analisis korelasi kanonik dengan menggunakan metode *robust* adalah analisis korelasi kanonik yang tidak terlalu dipengaruhi oleh pencilan.

Salah satu metode *robust* yang memiliki kemampuan mengukur jarak sekaligus mendeteksi pencilan adalah *Minimum Covariance Determinant* (MCD) yang diperkenalkan oleh Rousseeuw (1984). MCD adalah suatu metode yang menaksir rata-rata dan variansi-kovariansi dengan mencari himpunan bagian dari data yang menghasilkan determinan matriks variansi kovariansi terkecil.

Suatu ukuran tingkat ketaksensitifan suatu metode *robust* terhadap keberadaan pencilan digunakan *breakdown point*. *Breakdown point* adalah jumlah pengamatan minimal yang dapat menggantikan sejumlah pengamatan awal yang berakibat pada nilai taksiran yang dihasilkan sangat berbeda dari taksiran sebenarnya. Dengan *breakdown point* dapat diketahui seberapa banyak pencilan dalam data yang dapat membuat penaksir berubah. Menurut Hubert (2009), MCD tergolong penaksir *robust* yang baik karena memenuhi batas jumlah maksimum pencilan dalam data, yaitu sebesar 50%.

Berdasarkan uraian di atas, penulis tertarik untuk mengkaji *minimum covariance determinant* sebagai salah satu metode *robust* dalam suatu karya tulis, selanjutnya tugas akhir ini penulis beri judul “ANALISIS KORELASI KANONIK *ROBUST* DENGAN METODE *MINIMUM COVARIANCE DETERMINANT* (MCD)”.

1.2. Rumusan Masalah

Berdasarkan uraian di atas, maka permasalahan dalam penulisan tugas akhir ini dirumuskan sebagai berikut:

1. Bagaimanakah menentukan korelasi kanonik dengan metode MCD?
2. Bagaimanakah penerapan metode MCD dalam analisis korelasi kanonik *robust* dalam data pada studi kasus angkutan umum?

1.3. Batasan Masalah

Agar pembahasan masalah lebih spesifik pada penulisan tugas akhir ini dibahas untuk $p + q$ variabel dengan n pengamatan di mana $p + q < n$. Matriks variansi-kovariansi dari MCD adalah *non singular* dan tanpa membedakan jenis pencilan.

1.4. Tujuan Penulisan

Tujuan dari penulisan tugas akhir ini adalah :

1. Untuk mengetahui bagaimana menentukan koefisien korelasi kanonik dengan metode MCD.
2. Untuk mengetahui penerapan metode MCD dalam analisis korelasi kanonik *robust* dalam data pada studi kasus.

1.5. Manfaat Penulisan

1. Teoritis

Analisis korelasi kanonik *robust* dengan metode MCD adalah salah satu teknik analisis statistik, yang digunakan untuk melihat hubungan antara himpunan variabel bebas (X_1, X_2, \dots, X_p) dengan himpunan variabel terikat (Y_1, Y_2, \dots, Y_q) dan memiliki kekokohan terhadap pencilan. Metode MCD memberikan estimasi

dan kecocokan model yang lebih baik dibandingkan dengan metode statistik klasik.

2. Praktis

Melalui tugas akhir ini diharapkan dapat memberikan pengetahuan tambahan mengenai analisis korelasi kanonik *robust* dengan metode *minimum covariance determinant* (MCD), sehingga apabila menghadapi kasus yang sesuai dengan pembahasan dalam tugas akhir ini, dapat dipraktekkan dalam menyelesaikan masalah tersebut.

1.6. Sistematika Penulisan

Adapun sistematika penulisan pada tugas akhir ini adalah:

- BAB I** Pendahuluan
Mengemukakan latar belakang masalah, rumusan masalah, tujuan penulisan, manfaat penulisan dan sistematika penulisan.
- BAB II** Landasan Teori
Mengemukakan landasan teori yang mendukung Bab III di antaranya menjelaskan koefisien korelasi kanonik, *canonical loadings* dan *canonical cross loadings*.
- BAB III** Analisis Korelasi Kanonik Robust Dengan Metode *Minimum Covariance Determinant*
Membahas tentang analisis korelasi kanonik *robust* dengan metode *Minimum covariance determinant* (MCD).
- BAB IV** Studi Kasus

Pembahasan studi kasus penggunaan dari bahasan dalam Bab III, studi kasus yang diambil oleh penulis adalah data kinerja pelayanan operator dan evaluasi pelayanan angkutan umum.

BAB V Kesimpulan dan Saran

Mencoba merangkum keseluruhan hasil pembahasan dalam bentuk kesimpulan dan saran.

DAFTAR PUSTAKA

LAMPIRAN

