

Workshop Analisis Faktor untuk Data Penelitian Ilmu Sosial dan Kependidikan

Factor Analysis Workshop for Social and Educational Research Data

Dhoriva Urwatul Wutsqa, Kismiantini, Muhammad Fauzan, Rosita Kusumawati, Sahid dan Syarifah Inayati

Jurusan Pendidikan Matematika FMIPA UNY, Email: dhoriva_uw@uny.ac.id

Abstrak

Kegiatan PPM ini bertujuan memberikan pelatihan penggunaan program R dan SPSS dalam penelitian data ilmu sosial dan kependidikan kepada para praktisi lulusan S1, mahasiswa pasca sarjana, dan dosen di lingkungan universitas di Yogyakarta. Penelitian dalam ilmu sosial dan kependidikan seringkali melibatkan banyak variabel yang saling berkorelasi maupun mempunyai korelasi yang tinggi, misalnya dalam pengembangan suatu instrument penelitian. Penghapusan variabel yang mempunyai korelasi tinggi bisa mengakibatkan hilangnya informasi, untuk mengatasi hal tersebut dapat digunakan analisis komponen utama (*Principal Component Analysis/PCA*). Sedangkan dalam hal menentukan konstruk yang sesuai dari butir-butir soal yang terbentuk dapat dilakukan dengan dua pendekatan yaitu analisis faktor eksplorasi (*Exploratory Factor Analysis/EFA*) dan analisis faktor konfirmasi (*Confirmatory Factor Analysis/CFA*). Permasalahan yang terkait dengan reduksi variabel dan pembentukan konstruk pada pengembangan instrument ini merupakan hal yang sangat penting dalam penelitian ilmu sosial dan kependidikan. Sehingga suatu metode analisis PCA, EFA, dan CFA dengan menggunakan program R dan SPSS mutlak diperlukan. Workshop yang diikuti oleh 26 peserta ini dapat berjalan dengan baik. Berdasarkan pelaksanaan dan evaluasi kegiatan ini dapat disimpulkan bahwa tujuan dapat tercapai dengan baik.

Kata kunci: analisis faktor, analisis komponen utama, analisis faktor eksplorasi, analisis faktor konfirmasi

Abstract

This community service program aims to train undergraduate and master graduates, and lecturers at various universities in Yogyakarta to be able to use R and SPSS in social and educational research data. Research in social science and education involves many variables that are correlated or have a high correlation such as in developing a research instrument. Removing variables that have a high correlation can cause loss of information. To overcome this problem, principal component analysis (PCA) can be used. Meanwhile, in terms of determining the appropriate construction from the items that can be done with a two approach, namely exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). Problems related to variable reduction and making instruments for developing this instrument are very important in social science and educational research data. So that the use of PCA, EFA, and CFA in R and SPSS is required. The workshop which was attended by 26 participants went well. Based on these implementation and activities, it can be denied that achieving the goals can be achieved well.

Key words: *factor analysis, principal component analysis, exploratory factor analysis, confirmatory factor analysis*

PENDAHULUAN

Penelitian dalam ilmu sosial dan kependidikan seringkali melibatkan banyak variabel yang mempunyai korelasi yang tinggi. Jika akan dilakukan analisis regresi, korelasi yang tinggi ini menjadi salah satu permasalahan pelanggaran asumsi. Penghapusan variabel yang mempunyai korelasi yang tinggi bisa mengakibatkan hilangnya informasi. Hal tersebut dapat diatasi dengan analisis komponen utama (*Principal Component Analysis/PCA*). Tujuan dari analisis komponen utama adalah mereduksi variabel dengan membentuk komponen baru yang tidak berkorelasi, yang melibatkan semua variabel. Dengan demikian informasi variabel-variabel awal tidak hilang.

Pengembangan instrument dalam penelitian ilmu sosial dan kependidikan biasanya juga melibatkan banyak butir/item yang saling berkorelasi. Butir-butir ini dalam analisis selanjutnya dipandang sebagai variabel. Untuk menentukan konstruk yang sesuai dari butir-butir yang terbentuk dapat dilakukan dengan dua pendekatan. Jika konstruk dibentuk secara eksplorasi berdasarkan data empiris, maka dapat dilakukan dengan pendekatan analisis factor eksplorasi (*Exploratory Factor Analysis/EFA*). Dalam metode ini banyaknya faktor yang terbentuk, maupun pengelompokan variabel (struktur faktor) tidak ditetapkan terlebih dahulu, tetapi menggunakan kriteria secara statistik tertentu berdasarkan hasil data empiris. Pendekatan yang kedua adalah dengan analisis factor konfirmasi (*Confirmatory Factor Analysis/CFA*). Metode ini biasanya berdasarkan pada teori yang kuat. Jadi struktur dan banyak faktor sudah ditentukan terlebih dahulu berdasarkan kejian teori yang kuat. Analisis bertujuan untuk menentukan apakah struktur yang telah ditetapkan didukung oleh data empiris.

Permasalahan yang terkait dengan reduksi variabel dan pembentukan konstruk pada pengembangan instrument merupakan hal yang sangat penting dalam penelitian

ilmu sosial dan kependidikan. Para praktisi, dosen, calon dosen di bidang tersebut perlu memahami dengan baik metode analisis PCA, EFA, dan CFA. Selain itu, mereka juga harus trampil dalam mengaplikasikan program untuk analisis tersebut, serta dapat menginterpretasikan hasil analisis dengan benar.

Prodi Statistika di UNY mempunyai komitmen untuk mengembangkan kurikulum yang salah satunya adalah dengan membentuk konsentrasi statistika untuk ilmu sosial dan kependidikan. Sebagai salah satu bentuk komitmen, prodi menyelenggarakan kegiatan pelatihan atau workshop analisis data statistik untuk penelitian di bidang ilmu sosial dan kependidikan. Oleh karena itu dalam kegiatan pengabdian ini diusulkan workshop analisis faktor untuk data penelitian ilmu sosial dan kependidikan.

SOLUSI/TEKNOLOGI

Dalam rangka memberikan pelatihan tentang analisis data penelitian ilmu sosial dan kependidikan ini, selain keterampilan menyelesaikan masalah menggunakan komputasi statistik menggunakan program R dan SPSS, pemahaman konsep secara teoritis juga sangat diperlukan. Berikut pembahasan secara terkait PCA, EFA, dan CFA.

Analisis Komponen Utama (*Principal Component Analysis/PCA*)

PCA adalah prosedur matematis yang mengubah sekumpulan variabel yang berkorelasi menjadi sekumpulan variabel tidak berkorelasi yang disebut komponen utama, yaitu kombinasi linear dari variabel asal dan dapat dianggap sebagai variabel "baru" (Johnson & Wichern, 2007).

Kegunaan PCA meliputi: (1) penyaringan data, (2) *clustering*, (3) klasifikasi prediksi, (4) analisis regresi. Pada penyaringan data, PCA digunakan untuk melihat data dengan jumlah dimensi yang lebih kecil serta untuk mengidentifikasi pencilaan. Pada *clustering*, PCA digunakan

untuk menentukan bagaimana mengelompokan pengamatan. Pada klasifikasi prediksi, PCA digunakan untuk klasifikasi prediksi bila klasifikasi dari pengamatan sudah diketahui. Pada analisis regresi, PCA digunakan untuk menampilkan variabel-variabel independen yang berkorelasi tinggi menjadi sekumpulan variabel yang lebih kecil dimensinya.

Tujuan dari PCA adalah menemukan dimensi sebenarnya dari data tanpa menghilangkan informasi, dan menginterpretasi komponen utama (variabel baru). Sedangkan karakteristik dari komponen utama adalah tidak berkorelasi, komponen utama pertama menjelaskan sebanyak mungkin variabilitas dalam data, dan setiap komponen utama berturut-turut menjelaskan sebanyak mungkin variabilitas dalam data. Keunggulan dari PCA adalah tidak memerlukan asumsi distribusi normal.

PCA dapat dilakukan berdasarkan dengan matriks kovarians atau matriks korelasi. Pada PCA dengan matriks kovarians, sekumpulan variabel-variabel yaitu $\mathbf{x} \sim (\boldsymbol{\mu}, \boldsymbol{\Sigma})$ dengan \mathbf{x} memiliki dimensi $p \times 1$. Komponen utama pertama dinyatakan sebagai berikut: $y_1 = \mathbf{a}'_1(\mathbf{x} - \boldsymbol{\mu}) = a_{11}(x_1 - \mu_1) + a_{21}(x_2 - \mu_2) + \dots + a_{p1}(x_p - \mu_p)$ dengan \mathbf{a}_1 adalah vektor $p \times 1$ terpilih sehingga $Var[\mathbf{a}'_1(\mathbf{x} - \boldsymbol{\mu})]$ dimaksimalkan oleh semua vektor \mathbf{a}_1 dengan panjang 1. Nilai maksimum dari variansi adalah nilai eigen terbesar dari $\boldsymbol{\Sigma}$ yaitu λ_1 . Variansi tersebut ada bila \mathbf{a}_1 adalah vektor eigen yang bersesuaian dengan $\boldsymbol{\Sigma}$. Oleh karena variansi dari $y_1 = \mathbf{a}'_1(\mathbf{x} - \boldsymbol{\mu})$ dimaksimalkan, variabel baru y_1 akan menjelaskan sebanyak mungkin variabilitas dari \mathbf{x} . Komponen utama kedua adalah $y_2 = \mathbf{a}'_2(\mathbf{x} - \boldsymbol{\mu})$ dengan \mathbf{a}_2 terpilih sehingga $Var[\mathbf{a}'_2(\mathbf{x} - \boldsymbol{\mu})]$ dimaksimalkan oleh semua vektor \mathbf{a}_2 dengan panjang 1, serta y_1 dan y_2 tidak berkorelasi. Tidak berkorelasi berakibat bahwa \mathbf{a}_1 dan \mathbf{a}_2 adalah orthogonal ($\mathbf{a}'_1\mathbf{a}_2 = 0$). Nilai maksimum dari λ_2 (Nilai eigen terbesar kedua dari $\boldsymbol{\Sigma}$) dan muncul jika \mathbf{a}_2 adalah vektor eigen yang bersesuaian dengan $\boldsymbol{\Sigma}$.

Komponen utama ketiga, keempat, dan seterusnya adalah serupa langkahnya serupa dengan komponen utama pertama dan kedua. Jika vektor-vektor eigen orthogonal maka komponen utama juga orthogonal. Selanjutnya cara-cara menentukan jumlah komponen utama adalah (1) cari nilai terkecil d sehingga $(\sum_{j=1}^d \lambda_j / \sum_{i=1}^p \lambda_i) > 1$, di mana d adalah jumlah komponen utama yang terpilih, (2) dari *scree plot*, yaitu plot $\lambda_1, \lambda_2, \dots, \lambda_p$ vs. $1, 2, \dots, p$ dengan kriteria nilai d diperoleh dari jumlah nilai eigen sebelum level plot mendekati nol.

Permasalahan dalam PCA berdasarkan matriks kovarians adalah (1) jika sekumpulan variabel asal sudah tidak berkorelasi maka PCA tidak akan membantu, (2) PCA umumnya tidak menghilangkan variabel karena komponen utama adalah kombinasi linier dari variabel asal, dan (3) variabel asal perlu diukur dalam satuan yang sama dan memiliki variansi yang sama. Sehingga sebagai alternatif adalah menggunakan matriks korelasi \mathbf{R} . Cara penentuan jumlah komponen utama dengan PCA dari matriks korelasi serupa pada PCA dengan matriks kovarians.

Analisis Faktor Eksploratori (*Exploratory Factor Analysis/EFA*)

Pada berbagai penelitian dalam bidang psikologi dan disiplin ilmu perilaku lainnya, seringkali variabel yang diminati tidak dapat diukur secara langsung, seperti kecerdasan dan kelas sosial. Dalam kasus ini, peneliti dapat mengumpulkan informasi dari variabel terukur atau teramati secara langsung sebagai indikator untuk mengukur variabel yang diminati namun tak terukur secara langsung. Misal seorang psikolog ingin mengetahui kecerdasan individu melalui variabel teramati seperti nilai ujian dari berbagai macam disiplin ilmu. Misal seorang sosiolog ingin mengetahui kelas sosial masyarakat dengan mengajukan pertanyaan tentang pekerjaan, latar belakang pendidikan, kepemilikan rumah, dan lain-lain dari masyarakat tersebut.

Variabel kecerdasan dan kelas sosial ini yang disebut sebagai variabel laten (yaitu, variabel tak terukur secara langsung namun dapat diperoleh dari sejumlah variabel terukur). Metode analisis yang paling umum digunakan untuk mengetahui hubungan antara variabel laten dan variabel terukur (variabel *manifest*) adalah analisis faktor. Model yang menjadi dasar metode ini adalah regresi berganda, dengan variabel terukur sebagai variabel penjelas dan variabel laten yang tidak teramati secara langsung (disebut sebagai faktor umum). Oleh karena variabel laten tidak diketahui sebelumnya maka estimasi langsung terhadap koefisien regresi (*load factor*) tidak dapat dilakukan.

Ada dua tipe analisis faktor yaitu analisis faktor eksploratori (EFA) yang digunakan untuk menyelidiki hubungan antara variabel dan faktor manifestasi tanpa mengasumsikan variabel *manifest* mana yang terkait dengan faktor, dan analisis faktor konfirmatori (AFE) yang digunakan untuk menguji apakah model faktor spesifik yang dipostulasikan secara apriori cukup sesuai untuk kovarians atau korelasi antara variabel terukur. Pada bagian ini, akan dibahas AFE yang diperkenalkan oleh Spearman (1904) dengan model satu faktor.

Contoh sederhana dari model analisis faktor yaitu Spearman tertarik mengetahui kecerdasan anak berdasarkan pada tiga mata pelajaran, Klasik (X_1), Prancis (X_2), dan bahasa Inggris (X_3). Jika diasumsikan satu faktor tunggal maka model analisis faktor adalah

$$\begin{aligned} X_1 &= \lambda_1 F + \varepsilon_1 \\ X_2 &= \lambda_2 F + \varepsilon_2 \\ X_3 &= \lambda_3 F + \varepsilon_3 \end{aligned}$$

Model ini pada dasarnya adalah model regresi linier sederhana dari masing-masing variabel yang diamati pada satu faktor umum. Dalam contoh ini, variabel laten atau faktor umum, F , disebut sebagai kecerdasan atau kemampuan intelektual umum. Notasi $\lambda_1, \lambda_2, \lambda_3$ adalah koefisien regresi, dalam konteks ini, disebut sebagai *factor loadings*, dan $\varepsilon_1, \varepsilon_2, \varepsilon_3$ menyatakan kesalahan acak (*random error*). Kesalahan acak akan

memiliki variansi kecil jika variabel *manifest* terkait erat dengan variabel laten.

Model yang mendasari analisis faktor adalah model regresi yang menghubungkan variabel *manifest* dengan sekumpulan variabel laten yang tidak teramati. Model ini mengasumsikan bahwa hubungan yang diamati antara variabel *manifest* (diukur oleh matriks kovarians atau matriks korelasi) adalah hasil dari hubungan variabel-variabel ini dengan variabel laten. Oleh karena matriks kovarians atau korelasi dari variabel *manifest* adalah pusat bagi analisis faktor, maka diasumsikan bahwa variabel *manifest* memiliki rata-rata nol. Misal ada sebanyak p variabel *manifest* (X_1, X_2, \dots, X_p) dan k variabel laten dengan k faktor umum (F_1, F_2, \dots, F_k) dengan $k < p$, maka model analisis k faktor adalah

$$\begin{aligned} X_1 &= \lambda_{11}F_1 + \lambda_{12}F_2 + \dots + \lambda_{1k}F_k + \varepsilon_1 \\ X_2 &= \lambda_{21}F_1 + \lambda_{22}F_2 + \dots + \lambda_{2k}F_k + \varepsilon_2 \\ &\vdots \end{aligned}$$

$X_p = \lambda_{p1}F_1 + \lambda_{p2}F_2 + \dots + \lambda_{pk}F_k + \varepsilon_p$
 λ_{ij} adalah koefisien regresi dari variabel-variabel X pada faktor utama, dalam konteks analisis faktor, koefisien regresi ini disebut sebagai *factor loadings* yang menunjukkan sejauhmana variabel teramati X_i bergantung terhadap faktor utama. Faktor loading ini digunakan dalam interpretasi faktor. Persamaan regresi di atas dapat ditulis lebih ringkas sebagai berikut

$$\mathbf{X} = \mathbf{\Lambda F} + \boldsymbol{\varepsilon}$$

dengan

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_{11} & \dots & \lambda_{1k} \\ \vdots & & \vdots \\ \lambda_{p1} & \dots & \lambda_{pk} \end{pmatrix}, \mathbf{F} = \begin{pmatrix} F_1 \\ \vdots \\ F_k \end{pmatrix}, \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_p \end{pmatrix}.$$

Kesalahan acak $\varepsilon_1, \dots, \varepsilon_p$ diasumsikan tidak berkorelasi dengan satu sama lain dan dengan faktor F_1, \dots, F_k . Elemen-elemen dari $\boldsymbol{\varepsilon}$ adalah spesifik untuk masing-masing X_i dan disebut sebagai variat spesifik. Kedua asumsi tersebut berarti bahwa nilai-nilai dari faktor umum dan variabel *manifest* independen dan korelasi variabel teramati muncul dari hubungan

variabel tersebut dengan faktor umum. Oleh karena faktor-faktor tersebut tidak teramati, seringkali diasumsikan memiliki rata-rata nol dan simpangan baku satu. Asumsi lainnya adalah faktor-faktor tersebut tidak berkorelasi satu sama lain. Faktor loading adalah korelasi dari variabel *manifest* dan faktor-faktornya. Sehingga variansi dari variabel X_i adalah

$$\sigma_i^2 = \sum_{j=1}^k \lambda_{ij}^2 + \psi_i$$

dengan ψ_i adalah variansi dari ε_i . Dalam model analisis faktor ini, variansi dari masing-masing variabel teramati dapat dibagi menjadi dua bagian, yang pertama $h_i^2 = \sum_{j=1}^k \lambda_{ij}^2$ disebut sebagai komunalitas dari variabel yang menyatakan variansi dibagikan dengan variabel lain melalui faktor umum. Bagian kedua, ψ_i , disebut sebagai variansi spesifik atau unik dan menyatakan keragaman dari X_i yang tidak dibagi dengan variabel lainnya. Model faktor yang mengikuti bentuk kovariansi dari X_i dan X_j adalah:

$$\sigma_{ij} = \sum_{l=1}^k \lambda_{il} \lambda_{jl}$$

Kovariansi ini tidak tergantung dengan variat spesifik, hanya faktor umum saja yang bertujuan untuk menjelaskan hubungan antar variabel *manifest*. Model analisis k faktor dengan matriks kovariansi populasi (Σ) dapat dituliskan sebagai berikut

$$\Sigma = \Lambda \Lambda' + \Psi$$

dengan $\Psi = \text{diag}(\psi_i)$. Secara praktis, Σ akan diestimasi oleh matriks kovariansi sampel S dan matriks kovariansi sampel ini akan digunakan untuk mengestimasi Λ dan Ψ . Selanjutnya ingin ditentukan pula nilai k (banyaknya faktor).

Pendugaan dalam analisis faktor bertujuan untuk memperoleh $\hat{\Lambda}$ (matriks faktor loading dugaan) dan $\hat{\Psi}$ (matriks diagonal yang memuat variansi spesifik dugaan) dari matriks kovariansi sampel S .

$$S \approx \hat{\Lambda} \hat{\Lambda}' + \hat{\Psi}$$

Dari matriks faktor loading dugaan $\hat{\Lambda}$, maka variansi spesifik dugaan adalah

$$\hat{\psi}_i = s_i^2 - \sum_{j=1}^k \hat{\lambda}_{ij}^2, i = 1, \dots, p.$$

Ada dua metode pendugaan estimasi pada analisis faktor yaitu analisis faktor utama (*principal factor analysis*) dan analisis faktor likelihood maksimum (*maximum likelihood factor analysis*). Analisis faktor utama adalah teknik nilai eigen dan vektor eigen yang mirip dengan analisis komponen utama tetapi pengoperasian tidak secara langsung pada S (atau R), yaitu pada matriks kovariansi tereduksi, S^* didefinisikan sebagai

$$S^* = S - \hat{\Psi}$$

dengan $\hat{\Psi}$ adalah matriks diagonal yang memuat estimator ψ_i .

Analisis faktor likelihood maksimum mengasumsikan bahwa data mengikuti distribusi normal multivariat. Berdasarkan asumsi ini dan asumsi model analisis faktor berlaku, fungsi kemungkinan L dapat ditunjukkan sebagai $-\frac{1}{2}nF$ ditambah fungsi dari pengamatan dengan F yang didefinisikan sebagai

$$F = \ln|\Lambda \Lambda' + \Psi| + \text{trace}(S|\Lambda \Lambda' + \Psi|^{-1}) - \ln|S| - p$$

Fungsi F bernilai nol jika $\Lambda \Lambda' + \Psi$ sama dengan S dan nilai lebih besar dari nol jika tidak. Estimasi *loading* dan variansi spesifik diperoleh dengan meminimalkan F terhadap parameter-parameter ini. Nilai awal dari faktor loading dan variasi spesifik dapat diperoleh dalam beberapa cara. Pemilihan banyaknya faktor k dapat dilakukan dengan memeriksa solusi yang sesuai dengan nilai k berbeda dan menentukan secara subjektif untuk memperoleh interpretasi yang paling meyakinkan, atau dengan pendekatan diagram *scree*.

Pada faktor loading seringkali ditemui kendala dalam memperoleh estimator parameter dalam model analisis faktor yang unik. Solusi orthogonal lebih mudah diinterpretasikan dengan penggunaan model dengan loading $\Lambda^* = \Lambda M$, dengan M disebut sebagai rotasi faktor. Rotasi faktor digunakan supaya model analisis faktor dapat dideskripsikan sesederhana mungkin; rotasi ini tidak mengubah keseluruhan struktur solusi. Jadi rotasi adalah proses

supaya memperoleh solusi yang lebih mudah ditafsirkan tanpa mengubah sifat matematisnya. Ada dua tipe rotasi yaitu rotasi orthogonal (metode ini mensyaratkan bahwa faktor yang dirotasi itu tidak berkorelasi) dan rotasi *oblique* (metode ini memperbolehkan faktor-faktor berkorelasi). Rotasi ortogonal diperoleh dengan mengalikan matriks loading asal dengan matriks orthogonal. Rotasi *oblique*, matriks loading asal dikalikan dengan matriks yang tidak lagi dibatasi harus orthogonal. Jika rotasi orthogonal digunakan maka matriks korelasi antara faktor setelah rotasi berupa matriks identitas. Jika rotasi *oblique* digunakan maka matriks korelasi akan diharuskan memiliki elemen unit pada diagonalnya, namun tidak ada batasan pada elemen off-diagonal.

Ada dua teknik yang sering digunakan dalam rotasi orthogonal yaitu varimax dan quartimax. Rotasi varimax, diusulkan oleh Kaiser (1958), yaitu faktor-faktor dengan beberapa loading besar dan loading sedapat mungkin mendekati nol. Hal ini dicapai dengan memaksimalkan maksimal fungsi kuadrat dari loading yang diberikan oleh Mardia et al. (1979). Hal ini menghasilkan faktor-faktor yang memiliki korelasi tinggi dengan satu kumpulan variabel-variabel dan sedikit atau tidak ada korelasi dengan kumpulan lainnya. Rotasi quartimax, disarankan oleh Carroll (1953), mengharuskan suatu variabel yang diberikan untuk berkorelasi tinggi dengan satu faktor dan sama sekali tidak atau sangat rendah pada faktor lainnya. Rotasi quartimax tidak sepopuler daripada varimax.

Pada rotasi *oblique*, ada dua metode yang sering digunakan yaitu oblimin dan promax. Rotasi oblimin, ditemukan oleh Jennrich dan Sampson (1966), untuk menemukan struktur sederhana dari matriks pola faktor melalui parameter yang digunakan untuk mengendalikan tingkat korelasi antara faktor-faktor tersebut. Rotasi promax, diperkenalkan oleh Hendrickson dan White (1964), yaitu dengan menaikkan nilai loading dalam solusi ortogonal (umumnya rotasi varimax). Tujuannya

adalah untuk memperoleh solusi dengan struktur terbaik tetapi dengan loading serendah mungkin dan korelasi antara faktor-faktor juga terendah.

Analisis Faktor Konfirmatori (Confirmatory Factor Analysis/CFA)

CFA secara khusus menentukan suatu solusi dari model dan selanjutnya merotasi *factor loading* untuk memperoleh faktor laten yang lebih mudah diinterpretasikan. Sedangkan analisis faktor konfirmatori digunakan untuk merepresentasikan variabel laten yang menjadi perhatian namun sulit atau tidak mungkin diukur secara langsung. Sebagai contoh, peneliti pendidikan tertarik mengetahui perilaku siswa tentang aspek-aspek pengalaman kelas mereka, dokter ingin memeriksa berbagai dimensi kualitas hidup para pasien kanker yang memperoleh perawatan.

Pada EFA lebih diutamakan untuk menentukan berapa banyak faktor yang diperlukan untuk menjelaskan hubungan dan menentukan struktur faktor. Sedangkan pada CFA adalah mengkonfirmasi keberadaan struktur faktor dari model yang sudah ada, dan menguji hipotesis apakah model konsisten dengan data pengamatan. Karakteristik dari CFA sederhana adalah masing-masing indikator adalah variabel kontinu dengan dua penyebab (faktor/variabel laten dan galat), galat pengukuran saling independen serta saling independen terhadap faktor, dan hubungan antara faktor tidak dianalisis (kovarians). Model CFA didefinisikan sebagai berikut:

$$\mathbf{X} = \boldsymbol{\mu} + \boldsymbol{\Lambda}\mathbf{F} + \boldsymbol{\varepsilon}$$

Pada model CFA, $\boldsymbol{\mu}$ seringkali diasumsikan sebagai $\mathbf{0}$, sedangkan pada CFA tidak perlu diasumsikan seperti pada EFA. Metode yang umum digunakan untuk mengestimasi parameter model CFA adalah metode likelihood maksimum (ML). ML mengestimasi parameter-parameter model (*factor loadings*, variansi, kovariansi, intersep, dan variansi galat) yang meminimumkan persamaan berikut:

$$F_{ML} = \frac{1}{2} \text{tr} \left[((\mathbf{S} - \boldsymbol{\Sigma})\boldsymbol{\Sigma}^{-1})^2 \right]$$

dengan \mathbf{S} adalah matriks kovariansi antar indikator-indikator teramati, dan $\mathbf{\Sigma}$ adalah matriks kovariansi antar indikator-indikator teramati. Algoritma ML berdasarkan asumsi bahwa indikator-indikator mengikuti distribusi normal multivariat. Jika asumsi normal multivariat tidak terpenuhi, alternatif metode estimasi parameter adalah *weighted least squares* (WLS) yaitu $F_{WLS} = (\mathbf{S} - \mathbf{\Sigma})' \mathbf{W}^{-1} (\mathbf{S} - \mathbf{\Sigma})$, dengan \mathbf{W} adalah matriks terboboti yang merupakan matriks kovarians asimptotik dari elemen-elemen yang terdapat dalam matriks kovarians sampel teramati.

Setelah parameter-parameter telah diestimasi, selanjutnya ingin mengetahui apakah model layak dengan data. Ada beberapa indeks pengepasan yang dapat digunakan untuk mengetahui kelayakan model yaitu *root mean square error of approximation* (RMSEA), *comparative fit index* (CFI) dan Tucker Lewis Index (TLI), *standardized root mean square* (SRMR), *Akaike information criteria* (AIC), *Bayesian information criteria* (BIC), dan *sample-size-adjusted BIC* (SBIC). Nilai RMSEA $\leq 0,05$ mengindikasikan kelayakan model yang baik, nilai $0,05 < \text{RMSEA} \leq 0,08$ menunjukkan kelayakan model yang cukup, RMSEA $> 0,08$ dikatakan model buruk (Kline, 2001). Nilai CFI dan TLI $\geq 0,95$ mengindikasikan model layak yang baik (Hu & Bentler, 1999). Nilai SRMR $\leq 0,08$ menunjukkan kelayakan model yang baik (Hu & Bentler, 1999). Nilai AIC, BIC, dan SBIC yang rendah menunjukkan model layak terhadap data.

Identifikasi model CFA merupakan hal yang penting untuk mengetahui apakah model teridentifikasi atau *over identified* atau *under identified*. Jika suatu model teridentifikasi maka ada informasi yang cukup untuk memperoleh sekumpulan estimator parameter yang unik. Beberapa petunjuk bagi identifikasi model adalah (1) variabel laten harus diskalakan dengan menentukan indikator marker atau menetapkan variansi dari faktor dengan nilai 1, (2) banyaknya informasi dalam matriks input (matriks varians dan kovarians) \geq

banyaknya parameter model yang diestimasi (*factor loadings*, variansi/kovariansi faktor, variansi/kovariansi galat indikator), (3) pada model satu faktor memerlukan minimum tiga indikator, (4) model dengan dua atau lebih faktor memerlukan minimal dua indikator per variabel laten (Brown, 2006).

HASIL DAN DISKUSI

Segala persiapan dilakukan untuk melaksanakan kegiatan PPM diantaranya: (1) Koordinasi persiapan seminar proposal PPM; (2) Menyusun bahan untuk seminar proposal PPM; (3) Koordinasi pelaksanaan PPM untuk menentukan peserta, hari, dan tanggal kegiatan PPM; (4) Mengirim undangan kepada peserta kegiatan PPM; dan (5) Merancang kegiatan workshop bagi peserta PPM.

Kegiatan workshop atau pelatihan ini dilaksanakan di Ruang Laboratorium Komputer lantai 3, Gedung laboratorium Jurusan Pendidikan Matematika. Pelatihan ini diikuti oleh 26 peserta yang terdiri dari para praktisi lulusan S1, mahasiswa pasca sarjana, dan beberapa dosen di lingkungan universitas di Yogyakarta. Pelatihan ini dilaksanakan pada tanggal 27 dan 28 Juni 2019 selama 16 jam yang terdiri atas kegiatan pemaparan materi dan tutorial selama lima sesi. *Sesi pertama* (3 jam), diisi dengan pemaparan materi terkait dengan analisis komponen utama dan praktek menggunakan Program R dan SPSS. *Sesi kedua* (3 jam), diisi dengan pemaparan materi terkait dengan analisis faktor eksploratori dan praktek menggunakan Program R dan SPSS. *Sesi ketiga* (3 jam), diisi dengan pemaparan materi terkait dengan analisis faktor konfirmatori dan praktek menggunakan Program R dan SPSS. *Sesi keempat* (3 jam), diisi dengan penerapan materi terkait dengan analisis faktor konfirmatori pada masalah instrument evaluasi dan praktek menggunakan Program R dan SPSS. Dan *sesi kelima* (4 jam), peserta workshop diberikan tugas mandiri untuk menyelesaikan masalah analisis data penelitian ilmu sosial dan kependidikan

menggunakan program R dan SPSS, kemudian peserta diminta untuk praktek dan presentasi hasil praktiknya. Pembicara dalam pelatihan ini adalah semua Tim PPM yang terdiri dari 6 dosen. Mahasiswa sebagai anggota dilibatkan dalam teknis pelaksanaan kegiatan. Metode yang digunakan dalam workshop adalah tutorial dan praktek secara langsung.

Dalam implementasinya, untuk mendukung kegiatan ini tim pengabdian membuat handout terkait materi workshop yang akan dilaksanakan. Selain materi terkait analisis faktor melalui PCA, CFA, dan EFA juga disertai implementasinya dengan menggunakan program R dan SPSS. Peserta juga diberikan handout terkait program R. Handout menggunakan program R ini dirancang dengan memberikan tutorial contoh penggunaan perintah dalam R dan beberapa soal latihan analisa data.

Pada awal pelaksanaan kegiatan pelatihan para peserta ditanya terlebih dahulu perangkat lunak apa saja yang sudah biasa digunakan untuk mengolah data statistika dan sejauh apa mengenal program R dan SPSS. Pada umumnya peserta belum pernah menggunakan program R, ada sebagian kecil dari peserta yang pernah mendengar mengenai program R tetapi belum pernah mempelajarinya. Sehingga agar saat penyampaian materi pelatihan dengan praktek menggunakan program R berjalan lancar, Dosen penyaji pada awal kegiatan harus membimbing dari dasar terkait penggunaan program R ini.

Dalam pelaksanaan kegiatan, peserta bersemangat mengikuti kegiatan pelatihan. Peserta dapat menggunakan perintah-perintah analisis faktor untuk data penelitian ilmu sosial dan kependidikan serta dapat memberikan interpretasi dari output yang diperoleh di SPSS dan R. Peserta terlihat cukup baik dari mulai pelatihan secara teoritis maupun workshop. Indikasi ini terlihat dari perhatian, pertanyaan yang mereka ajukan bila belum memahami, dan kesungguhan dalam praktek analisis data. Para peserta berusaha untuk menyelesaikan permasalahan yang diberikan hingga tuntas.

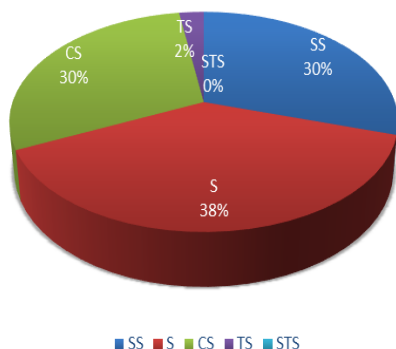
Interaksi juga terjalin antara peserta, mereka berdiskusi dengan teman terdekatnya untuk menyelesaikan tugas analisis data. Pada sesi terakhir, tugas yang dikerjakan secara diskusi kelompok pun terselesaikan dengan baik. Dengan demikian secara keseluruhan dapat disimpulkan bahwa acara workshop telah berjalan dengan lancar dan target acara telah tercapai.

Pada sesi terakhir, peserta diberikan angket evaluasi untuk memberikan penilaian mengenai workshop yang diikuti. Evaluasi pelaksanaan kegiatan pengabdian ini dilakukan dengan memberikan angket secara online untuk mengukur tingkat keberhasilan dari pelaksanaan kegiatan. Adapun aspek yang ditanyakan dalam butir angket adalah sebagai berikut:

- (1) Peserta workshop memahami cara menginstall R dan R studio *user interface*.
- (2) Peserta workshop memahami manfaat mempelajari perangkat lunak statistika.
- (3) Peserta workshop dapat melakukan *entry/input* data pada perangkat lunak statistika.
- (4) Peserta workshop mampu menyajikan statistika deskriptif dengan memanfaatkan perangkat lunak statistika.
- (5) Peserta workshop memahami kegunaan EFA, PCA, dan CFA.
- (6) Peserta workshop memahami dasar teori EFA, PCA, dan CFA.
- (7) Peserta workshop memahami prosedur analisis dari EFA, PCA, dan CFA.
- (8) Peserta workshop mampu menginterpretasi output/tampilan hasil analisis dari EFA, PCA, dan CFA.
- (9) Peserta workshop mengetahui manfaat Workshop Analisis Faktor untuk Data Penelitian Ilmu Sosial dan Kependidikan.

Selain hal tersebut di atas responden juga memberikan kritik, saran, dan masukan yang bersifat membangun pada Tim PPM untuk perbaikan pelaksanaan kegiatan-kegiatan PPM selanjutnya.

Angket evaluasi disusun menggunakan skala Likert yang terdiri dari 5 pilihan skala yang mempunyai gradasi dari Sangat Setuju (SS) hingga Sangat Tidak Setuju (STS). Secara keseluruhan, persentase pilihan jawaban SS (Sangat Setuju), S (Setuju), CS (Cukup Setuju), TS (Tidak Setuju), dan STS (Sangat Tidak Setuju) tergambar pada diagram lingkaran yang disajikan pada Gambar 1.



Gambar 1. Hasil angket evaluasi workshop analisis faktor untuk data penelitian ilmu sosial dan kependidikan

Dari hasil angket evaluasi dapat diketahui bahwa sekitar 68% peserta menjawab setuju atau sangat setuju untuk hampir seluruh pertanyaan dalam angket yang diberikan, hal ini berarti mayoritas memahami materi pelatihan dan dapat mengikuti pelatihan dengan baik. Meskipun 30% merasa cukup setuju dan sebagian kecil yaitu 2% masih merasa tidak setuju, hal ini menunjukkan bahwa beberapa materi pelatihan masih memerlukan penekanan dan pengulangan dalam penyampaian dan pemberian contoh-contoh yang lebih banyak serta pendampingan terhadap peserta juga diperlukan, karena beberapa peserta masih dapat dikatakan baru dalam mengenal program R. Hasil ini akan menjadi bahan evaluasi bagi tim PPM untuk perbaikan kedepannya dalam mengadakan pelatihan. Dengan demikian secara keseluruhan dapat disimpulkan bahwa acara workshop telah berjalan dengan lancar, dan target acara telah tercapai.

Dalam pelaksanaan kegiatan PPM ini, adapun hambatan yang ditemui yaitu

kurangnya pemahaman awal peserta dalam penggunaan program R. Sebagian besar peserta, dapat dikatakan baru dalam mengenal program R saat mengikuti pelatihan, sehingga pengenalan program R perlu dikenalkan dari dasar-dasar penggunaan program. Untuk mengatasi hal tersebut, tim dosen/penyaji memberikan materi pengenalan program R secara efektif terkait materi workshop, dan tim PPM lainnya secara aktif mendampingi peserta dalam praktek penggunaannya.

KESIMPULAN

Berdasarkan pelaksanaan dan evaluasi kegiatan pelatihan ini dapat disimpulkan bahwa melalui pelatihan ini para peserta mendapat pengetahuan dan ketrampilan tentang analisis faktor untuk data penelitian dan kependidikan, baik menggunakan SPSS maupun dengan R. Kegiatan pengabdian ini juga meningkatkan kesadaran penggunaan perangkat lunak tak berbayar untuk analisa data statistika. Tim PPM berhasil menyusun handout metode analisis PCA, EFA, dan CFA dengan menggunakan program R dan SPSS untuk analisa data statistika.

UCAPAN TERIMA KASIH

Pelaksanaan kegiatan ini dapat terlaksana karena kerjasama semua pihak. Kami mengucapkan terimakasih kepada kemenristekdikti atas hibah PkM.

PUSTAKA

- Brown, T.A. 2006. *Confirmatory factor analysis for applied research*. New York: The Guilford Press.
- Carroll, J. B. 1953. An analytical solution for approximating simple structure in factor analysis. *Psychometrika*, 18, 23-38.
- Hendrickson, A., & White, P. 1964. Promax: A quick method for rotation to oblique simple structure. *British Journal of*

- Mathematical and Statistical Psychology*, 17, 65-70.
- Hu, L., & Bentler, P. M. 1999. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55.
- Jennrich, R., & Sampson, P. 1966. Rotation for simple loadings. *Psychometrika*, 31, 313-323.
- Johnson, R., & Wichern, D. 2007. *Applied multivariate statistical analysis 6th edition*. New Jersey: Pearson Education.
- Kaiser, H. 1958. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23, 187-200.
- Kline, R. B. 2005. *Principles and practices of structural equation modeling* (2nd ed.). New York: Guilford Press.
- Mardia, K. V., Kent, J. T., & Bibby, J. M. 1979. *Multivariate Analysis*. London: Academic Press.
- Spearman, C. 1904. General intelligence objectively determined and measured. *The American Journal of Psychology*, 15, 201-293.