

BAB I

PENDAHULUAN

1.1 Latar Belakang

Ketersediaan data yang melimpah menuntut proses pengolahan dan penyajian informasi yang berkualitas guna mendukung pengambilan keputusan yang efektif (Rahutomo dkk., 2019). Tanpa penerapan metode statistika yang tepat, aliran data yang masif tersebut tidak akan mampu menghasilkan keputusan yang bermakna (Levpušček & Cukon, 2022). Sebagai ilmu yang berfokus pada pengumpulan, perangkuman, dan penyajian data, statistika berperan penting dalam mengubah data mentah menjadi informasi yang bermanfaat. Penguasaan ilmu ini tidak hanya meningkatkan kemampuan berpikir kritis dan logis (Lehman & Nisbett, 1990; Vanderstoep & Shaughnessy, 1997), tetapi juga menjadi kompetensi esensial di berbagai ranah akademik dan profesional (Maat & Rosli, 2016). Meskipun sering dianggap identik dengan bidang eksakta, statistika juga memiliki peran yang penting dalam ilmu sosial. Dalam konteks ini, statistika membekali individu dengan keterampilan penelitian yang komprehensif, mulai dari perumusan masalah dan hipotesis, analisis dan interpretasi data, hingga penyusunan laporan yang sistematis.

Mata kuliah terkait statistika bagi sebagian mahasiswa sering dipandang sebagai mata kuliah yang memicu kecemasan. Perasaan kecemasan atau ketakutan terhadap mata kuliah Statistika dinamakan juga sebagai *statistics anxiety* (Chew & Dillon, 2014). Zeidner (1991) menjelaskan *statistics anxiety* adalah kondisi kecemasan yang ditandai dengan kekhawatiran berlebih, pikiran yang mengganggu, disorganisasi mental, ketegangan, dan rangsangan fisiologis ketika seseorang dihadapkan pada materi atau permasalahan statistik. Hal ini sejalan dengan pandangan Onwuegbuzie (1997) yang mendefinisikan *statistics anxiety* sebagai kecemasan yang muncul saat mahasiswa berhadapan dengan statistik dalam bentuk dan tingkat kesulitan apa pun. Onwuegbuzie & Wilson (2003) menyatakan bahwa sekitar 75% mahasiswa pascasarjana mengalami *statistics anxiety*.

Penelitian tentang *statistics anxiety* pada mahasiswa telah banyak dilakukan dengan memanfaatkan instrumen pengukuran yang telah tersedia maupun yang dikembangkan secara khusus. Sebagai contoh, Iriyadi dkk. (2024) mengkaji tingkat kecemasan statistik pada mahasiswa non-eksakta sebelum dan setelah mengikuti mata kuliah Statistika dengan menggunakan instrumen *Statistical Anxiety Scale* (SAS) yang dikembangkan berdasarkan karakteristik pembelajaran di Indonesia. Hasil analisis menggunakan model Rasch dan CFA menunjukkan bahwa instrumen tersebut memiliki kinerja yang baik serta tidak terdapat butir pernyataan yang membingungkan responden dalam menjawab. Temuan penelitian Iriyadi dkk. (2024) juga menunjukkan adanya perubahan tingkat kecemasan mahasiswa setelah perkuliahan Statistika, di mana proporsi mahasiswa yang tergolong memiliki tingkat kecemasan rendah, tingkat kecemasannya meningkat dari 28,91% menjadi 34,16%. Sebaliknya, proporsi mahasiswa dengan tingkat kecemasan sedang mengalami penurunan, dari 37,57% menjadi 35,67%. Begitu pula proporsi mahasiswa dengan tingkat kecemasan tinggi mengalami penurunan dari 33,52% menjadi 30,17%. Hasil penelitian Iriyadi dkk. (2024) mengindikasikan bahwa partisipasi dalam perkuliahan statistika dapat memengaruhi tingkat kecemasan mahasiswa, terutama dalam konteks pembelajaran daring pada masa pandemi.

Levpušček & Cukon (2022) mengkaji hubungan antara *statistics anxiety*, *trait anxiety*, sikap terhadap konsep matematika dan statistika, serta pencapaian akademik mahasiswa. Levpušček & Cukon (2022) melibatkan mahasiswa yang telah menempuh setidaknya satu mata kuliah Statistika sebagai subjek penelitian. Pengukuran dilakukan menggunakan instrumen *Statistics Anxiety Rating Scale* (STARS), *State-Trait Anxiety Inventory*, serta kuesioner yang mengukur persepsi terhadap matematika dan statistika. Hasil penelitian Levpušček & Cukon (2022) menunjukkan bahwa skor rata-rata pada sebagian besar dimensi STARS tergolong rendah, kecuali pada dimensi *test anxiety* dan *class anxiety*. Selain itu, ditemukan bahwa mahasiswa perempuan memiliki tingkat kecemasan yang lebih tinggi dibandingkan mahasiswa laki-laki, serta tingkat *statistics anxiety* tertinggi terdapat pada kelompok mahasiswa yang memandang matematika dan statistika sebagai ancaman. Lebih lanjut,

rendahnya minat terhadap matematika sejak di bangku SMA menyebabkan mahasiswa memandang mata kuliah Statistika kurang bernilai, yang pada akhirnya berdampak pada rendahnya *computation self-concept*.

Instrumen yang telah digunakan dalam kajian *statistics anxiety*, termasuk SAS yang dikembangkan oleh Iriyadi dkk. (2024) dan STARS yang digunakan oleh Levpušček & Cukon (2022), pada dasarnya bekerja berdasarkan prinsip pengukuran konstruk laten. Secara umum, instrumen pengukuran seperti kuesioner merupakan alat yang digunakan peneliti untuk memperoleh informasi mengenai suatu konstruk, trait, atau dimensi tertentu. Ketiga hal tersebut merupakan entitas atau konsep teoretis (Desjardins & Bulut, 2018). Dalam konteks psikologi, sebuah konstruk umumnya bersifat laten sehingga tidak dapat diukur secara langsung. Konstruk tersebut diukur melalui variabel indikator (variabel manifes) yang dapat diamati seperti respons/jawaban peserta terhadap item pertanyaan kuesioner/tes (Bollen, 2002; Masyn dkk., 2010). Melalui sekumpulan variabel manifes konstruk laten dapat diukur secara kuantitatif.

Terdapat dua pendekatan metodologis utama untuk mengevaluasi suatu instrumen/alat ukur, yaitu *Classical Test Theory* (CTT) dan *Item Response Theory* (IRT). Kedua pendekatan tersebut mampu mengukur konstruk laten seperti *statistics anxiety* secara akurat. Teori Tes Klasik (*Classical Test Theory*/CTT) merupakan seperangkat teori psikometrik yang dikembangkan untuk memprediksi hasil pengukuran psikologis dengan mengasumsikan bahwa setiap item mencerminkan konstruk laten yang mendasarinya (Krabbe, 2017). Pendekatan CTT sangat bergantung pada konsep korelasi karena metode yang dominan pada masa tersebut adalah statistik korelasi Pearson (Baker & Kim, 2004). Menurut Hambleton dkk. (1991), kemampuan peserta dalam pendekatan CTT diartikan sebagai *true score*. Skor ini merupakan nilai harapan dari performa yang diamati sehingga maknanya selalu bergantung pada tes tertentu. Akibatnya, karakteristik peserta dan karakteristik tes menjadi saling bergantung dan tidak dapat dipisahkan. Parameter seperti tingkat kesulitan item, daya beda, reliabilitas, dan validitas sangat dipengaruhi oleh kelompok peserta yang digunakan untuk mengestimasi parameter tersebut. Ketergantungan inilah yang

menjadi kelemahan utama CTT dan mendorong penggunaan pendekatan yang lebih modern seperti *Item Response Theory* (IRT).

Teori Respons Item (*Item Response Theory*/IRT) yang juga dikenal sebagai *latent trait theory* pada dasarnya merupakan model yang merepresentasikan hubungan antara variabel laten dan manifestasinya (de Ayala, 2022; Raykov & Marcoulides, 2011). Pendekatan IRT dicirikan oleh asumsi bahwa konstruk yang dipelajari bersifat laten dan kontinu. Secara konseptual, prinsip dasar dalam IRT memandang bahwa respons peserta terhadap item instrumen dan performa tes secara keseluruhan dapat dijelaskan oleh satu atau lebih kemampuan laten yang jumlahnya jauh lebih sedikit dibandingkan total keseluruhan item (Raykov & Marcoulides, 2011). Pendekatan IRT memiliki dua asumsi utama, yaitu unidimensionalitas dan independensi lokal (Desjardins & Bulut, 2018). Asumsi unidimensionalitas menyatakan bahwa observasi terhadap variabel manifes semata-mata merupakan fungsi dari satu variabel laten individu yang kontinu (de Ayala, 2022). Sementara itu, asumsi independensi lokal menegaskan bahwa respons/jawaban seseorang dengan karakteristik atau sifat laten tertentu, terhadap suatu item tidak akan bergantung pada responsnya terhadap item lainnya (Desjardins & Bulut, 2018).

Dalam perkembangannya, salah satu model dalam kerangka IRT yang sering digunakan adalah Model Rasch (Nurcahyo, 2016). Model Rasch merupakan bentuk spesifik dari model *one-parameter logistic* (1PL), di mana seluruh item diasumsikan memiliki parameter daya pembeda yang seragam dan nilainya ditetapkan konstan sebesar satu (de Ayala, 2022; Desjardins & Bulut, 2018). Namun, Model Rasch standar memiliki keterbatasan karena mengasumsikan bahwa parameter item berlaku sama bagi seluruh responden. Dengan kata lain, model ini memandang responden sebagai bagian dari satu populasi yang relatif homogen, sehingga perbedaan pola respons antarindividu lebih banyak dijelaskan melalui perbedaan kemampuan atau tingkat trait laten responden, bukan melalui adanya perbedaan karakteristik kelompok dalam populasi (Rost, 1990).

Pada pengukuran konstruk afektif seperti *statistics anxiety*, asumsi homogenitas tersebut sering kali tidak terpenuhi karena individu dapat menunjukkan karakteristik respons yang berbeda meskipun berada pada tingkat karakteristik laten yang relatif sama. de Ayala & Santiago (2016) menyebutkan bahwa kondisi ini terjadi karena populasi dapat terdiri atas beberapa subpopulasi laten yang berbeda secara kualitatif dalam pola responsnya. De Meij dkk. (2008) juga menunjukkan bahwa individu dapat memperlihatkan pola berbeda dalam penggunaan kategori respons pada skala Likert. Pola ini tidak hanya mencerminkan konstruk yang diukur, melainkan juga kecenderungan gaya respons (*response styles*) tertentu, seperti preferensi kategori tengah atau respons ekstrem.

Untuk mengidentifikasi adanya perbedaan pola respons tersebut, salah satu pendekatan yang dapat digunakan adalah *Latent Class Analysis* (LCA). Berbeda dengan IRT yang menganggap variabel laten bersifat kontinu, dalam LCA variabel laten diasumsikan bersifat kategorik (de Ayala, 2022). LCA memungkinkan identifikasi kelas atau subkelompok tersembunyi dalam suatu populasi berdasarkan kombinasi fitur yang spesifik dan dapat diprediksi (Naldi & Cazzaniga, 2020). Pendekatan ini membagi subjek ke dalam kelas laten yang saling eksklusif atau tidak *overlap* dan mencakup keseluruhan (*mutually exclusive and exhaustive classes*) menggunakan variabel kategorik yang teramati (Collins & Lanza, 2010; O'Donnell dkk., 2017). Secara umum, penentuan jumlah kelas laten terbaik dilakukan dengan membandingkan beberapa model yang memiliki jumlah kelas berbeda. Sebagai contoh, membandingkan model satu kelas dengan model dua kelas untuk memilih struktur yang paling baik dalam menjelaskan data (de Ayala, 2022).

Meskipun LCA mampu mengidentifikasi subkelompok laten berdasarkan pola respons, pendekatan ini belum secara khusus memodelkan hubungan antara kemampuan atau trait laten individu dengan karakteristik item sebagaimana dilakukan dalam IRT. Oleh karena itu, dikembangkan pendekatan *Mixture Item Response Theory* (MixIRT), yang mengintegrasikan konsep kelas laten dalam LCA dengan pemodelan parameter item dalam IRT. MixIRT mengasumsikan keberadaan subpopulasi laten dalam data, sehingga parameter

item dan kemampuan dapat diestimasi secara spesifik pada setiap kelas laten/subpopulasi (Sen & Cohen, 2019). Dengan demikian, MixIRT tidak hanya mampu mengestimasi tingkat konstruk laten individu, tetapi juga mengelompokkan mereka berdasarkan pola respons secara probabilistik (Jiao dkk., 2011; Sen & Cohen, 2020).

Salah satu model MixIRT yang secara khusus diaplikasikan untuk data dikotomis adalah Model *Mixture Rasch* (*Mixture Rasch Model*/MRM). Model ini mengembangkan model Rasch standar melalui penambahan komponen kelas laten guna menangkap perbedaan struktur respons individu (Rost, 1990). Beberapa penelitian terkait efektivitas MRM telah dilakukan. Toker & Seidel (2023) menerapkan MRM dan menemukan dua kelas laten dengan pola jawaban berbeda. Mereka menyimpulkan bahwa instrumen berpotensi tidak mengukur konstruk yang sama pada seluruh kelompok sehingga dapat memengaruhi validitas. Widhiarso (2013) menunjukkan adanya heterogenitas dalam pola respons partisipan pada skala psikologis. Melalui analisis MRM, ditemukan tiga kelompok responden dengan karakteristik unik, termasuk kelompok yang tidak konsisten, yang jika diabaikan dapat mengganggu akurasi estimasi parameter item. Oleh karena itu, penggunaan MRM menjadi sangat penting untuk mengidentifikasi kelompok partisipan yang memiliki pola respons yang unik.

Penerapan pemodelan MRM tersebut menjadi sangat relevan ketika dihadapkan pada pengukuran instrumen psikologis. Dalam pengukuran *statistics anxiety* misalnya, salah satu alat ukur yang paling dikenal adalah *Statistical Anxiety Rating Scale* (STARS) yang dikembangkan oleh Cruise dkk. (1985). Instrumen terbagi ke dalam enam subskala, yakni *interpretation anxiety*, *test and class anxiety*, *fear of asking for help*, *worth of statistics*, *computation self concept*, dan *fear of statistics teachers*. Hanna dkk. (2008) menemukan adanya pola korelasi positif dan negatif yang bervariasi antarsubskala tersebut tergantung metode penskorannya. Temuan ini mengindikasikan bahwa beberapa subskala dalam STARS tidak mengukur *statistics anxiety* secara langsung, melainkan lebih mengarah pada konsep-konsep terkait seperti pandangan mengenai nilai atau kegunaan statistik.

Keterbatasan pada STARS kemudian mendorong pengembangan instrumen yang lebih spesifik dalam mengukur *statistics anxiety*. Vigil-Colet dkk. (2008) menyatakan bahwa hanya tiga subskala pertama pada STARS yang mengukur kecemasan secara langsung, sedangkan subskala lainnya lebih merepresentasikan persepsi mahasiswa terhadap statistika dan konsep diri. Sebagai penyempurnaan, Vigil-Colet dkk. (2008) mengembangkan *Statistical Anxiety Scale* (SAS) yang berfokus pada tiga dimensi utama, yaitu *interpretation anxiety*, *test and class anxiety*, dan *fear of asking for help*. Instrumen tersebut kemudian diperbarui oleh Lorenzo-Seva dkk. (2022) menjadi *Statistical Anxiety Scale-Revised* (SAS-R). Oleh karena itu, instrumen dalam penelitian ini disusun berdasarkan SAS-R yang dimodifikasi sesuai dengan konteks mahasiswa UNJ. Meskipun demikian, istilah SAS tetap digunakan sebagai istilah umum karena SAS-R merupakan versi revisi dari SAS dan tetap mengukur tiga dimensi utama *statistics anxiety*. Penggunaan instrumen ini juga diperkuat oleh penelitian Iriyadi dkk. (2024) yang menunjukkan bahwa adaptasi SAS dalam konteks Indonesia memiliki struktur dimensi yang jelas serta kualitas item yang baik melalui pendekatan Rasch dan CFA.

Berdasarkan rangkaian pemaparan tersebut, peneliti tertarik untuk mengidentifikasi keberadaan kelas laten *statistics anxiety* di kalangan mahasiswa. Penelitian ini berfokus pada mahasiswa program sarjana Universitas Negeri Jakarta (UNJ) angkatan 2022 hingga 2025 di Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) dan Fakultas Ilmu Sosial dan Hukum (FISH) yang telah menempuh mata kuliah Statistika. Pemilihan kedua fakultas tersebut tidak dimaksudkan untuk mewakili seluruh bidang eksakta dan non-eksakta, melainkan untuk menggambarkan dua konteks akademik yang memiliki karakteristik penggunaan statistika yang berbeda. Mahasiswa FMIPA memiliki paparan kuat terhadap konsep kuantitatif, sedangkan mahasiswa FISH cenderung menggunakan statistika sebagai alat analisis sosial. Perbedaan latar belakang dan pengalaman belajar ini relevan dikaji karena berkaitan dengan tingkat *statistics anxiety* yang dialami mahasiswa (Zeidner, 1991; Levpušček & Cukon, 2022). Melalui penggunaan

instrumen *Statistical Anxiety Scale* (SAS), respons mahasiswa terhadap setiap item akan dikotomisasi menjadi dua kategori (0 dan 1) sesuai kebutuhan analisis model *Mixture Rasch*. Pendekatan ini diharapkan mampu mengelompokkan mahasiswa secara probabilistik berdasarkan pola respons terhadap item-item *statistics anxiety* secara lebih spesifik.

1.2 Perumusan Masalah

Berdasarkan latar belakang, dapat dibuat rumusan masalah sebagai berikut:

1. Bagaimana kategorisasi kelas laten yang optimal dalam merepresentasikan heterogenitas *statistics anxiety* mahasiswa melalui pendekatan pemodelan *Mixture Rasch*?
2. Bagaimana karakteristik tingkat kesulitan item instrumen *statistics anxiety* pada setiap kelas laten?
3. Bagaimana dugaan tingkatan *statistics anxiety* dan pola respons mahasiswa pada masing-masing kelas laten tersebut?

1.3 Batasan Masalah

1. Penelitian ini dibatasi pada konstruk *statistics anxiety* mahasiswa secara menyeluruh yang diukur menggunakan 22 item dari *Statistical Anxiety Scale* (SAS). Analisis dilakukan secara simultan pada seluruh item sebagai satu kesatuan konstruk, tanpa memisahkan tiga dimensi teoretisnya (*examination anxiety*, *asking for help anxiety*, dan *interpretation anxiety*).
2. Respons mahasiswa terhadap item instrumen diperlakukan sebagai data dikotomis sesuai dengan kebutuhan analisis model *Mixture Rasch*.
3. Karakteristik item difokuskan pada estimasi tingkat kesulitan item yang bersifat spesifik pada masing-masing kelas laten. Selain itu, evaluasi kecocokan item dilakukan melalui statistik fit untuk mengetahui apakah item berfungsi secara konsisten.
4. Analisis hanya dilakukan pada item dan responden yang memiliki respons bervariasi. Item dengan jawaban seragam serta responden dengan pola jawaban seragam pada seluruh item dikeluarkan dari proses analisis.

5. Metode yang digunakan untuk menduga parameter dalam model *Mixture Rasch* adalah *Conditional Maximum Likelihood* (CMLE) dengan algoritma *Expectation-Maximization* (EM).
6. Pemodelan *Mixture Rasch* dibatasi pada satu hingga tiga kelas saja. Hal ini agar modelnya tetap sederhana, hasil hitungannya stabil, dan kelas yang terbentuk lebih mudah dipahami.

1.4 Tujuan Penelitian

1. Menganalisis kategorisasi kelas laten yang optimal dalam merepresentasikan heterogenitas *statistics anxiety* mahasiswa melalui pendekatan pemodelan *Mixture Rasch*.
2. Menganalisis karakteristik tingkat kesulitan item pada instrumen *statistics anxiety* untuk setiap kelas laten yang terbentuk.
3. Mengidentifikasi dan mendeskripsikan dugaan tingkat *statistics anxiety* serta pola respons mahasiswa pada masing-masing kelas laten yang terbentuk.

1.5 Manfaat Penelitian

Manfaat yang diperoleh dari penelitian ini adalah memberikan kontribusi bagi pengembangan keilmuan, khususnya dalam bidang statistika, psikometri, dan psikologi pendidikan. Penelitian ini diharapkan dapat memperkaya referensi mengenai penerapan MixIRT dalam menganalisis instrumen penelitian, terutama dalam mengungkap heterogenitas laten yang sering kali tidak terdeteksi oleh analisis statistik standar. Selain itu, secara praktis, hasil penelitian ini diharapkan menjadi informasi berharga bagi pihak program studi atau lembaga pendidikan untuk memetakan kelompok mahasiswa berdasarkan karakteristik *statistics anxiety* yang berbeda, sehingga dapat digunakan sebagai dasar dalam merancang intervensi, layanan bimbingan akademik, atau strategi pembelajaran yang lebih tepat.