

Perbandingan Penskoran Dikotomi dan Politomi dalam Teori Respon Butir untuk Pengembangan Bank Soal Matakuliah Matematika Dasar

Anisa¹

Abstrak

Teori Respon Butir merupakan pendekatan modern yang dilakukan untuk memberikan evaluasi dan penilaian terhadap perangkat tes yang digunakan dan hasil belajar peserta didik untuk mendapatkan informasi tentang sejauh mana keberhasilan guru/dosen sebagai fasilitator dalam menyajikan bahan pembelajaran serta sejauh mana peserta didik dapat menyerap materi yang diajarkan. Pada penelitian ini Teori Respon Butir difokuskan pada Penskoran Dikotomi dan Politomi untuk mengevaluasi perangkat tes matakuliah Matematika Dasar pada Fakultas Kehutanan dan Fakultas Ilmu Kelautan dan Perikanan Universitas Hasanuddin, dan penilaian hasil belajar mahasiswa pada kedua fakultas tersebut, sebagai langkah untuk pengembangan bank soal yang berkualitas untuk matakuliah ini. Hasil penelitian ini diharapkan dapat menjadi acuan untuk soal matakuliah Matematika Dasar yang merupakan matakuliah di bawah koordinasi dari UPT MKU Universitas Hasanuddin Makassar.

Kata Kunci:Teori Respon Butir, penskoran dikotomi, penskoran politomi.

1. Pendahuluan

Peranan pendidikan dalam kemajuan suatu bangsa dan masyarakat merupakan suatu keniscayaan. Pendidikan merupakan investasi jangka panjang yang harus selalu ditingkatkan mutunya [2]. Tujuan umum dari pendidikan adalah menjadikan peserta didik sebagai manusia yang mampu memecahkan permasalahannya dalam kehidupan sehari-hari. Salah satu proses penting dalam pendidikan adalah kegiatan penilaian atau evaluasi hasil belajar peserta didik yang harus dilakukan oleh pendidik, sebagai bekal bagi peserta didik untuk mampu beradaptasi dengan berbagai permasalahan yang bakal dihadapi nantinya setelah terjun ke masyarakat. Kemampuan memecahkan suatu permasalahan dapat dijadikan indikator mengukur ketercapaian dari tujuan pendidikan tersebut. Pelaksanaan evaluasi bertujuan untuk mengukur dan mengendalikan mutu pendidikan yang sedang dilaksanakan.

Hasil belajar adalah tingkat pencapaian belajar dari peserta didik [8]. Terdapat beberapa cara yang dapat dilakukan untuk mengetahui pencapaian hasil belajar peserta didik, salah satunya adalah melalui tes dengan menggunakan perangkat tes, baik tes pada ujian tengah semester, ujian akhir semester, maupun ujian akhir pada jenjang pendidikan di tingkat SD, SMP, SMU dan sederajatnya. Pemberian ujian atau tes dalam proses belajar mengajar dimaksudkan untuk mendapatkan informasi tentang sejauh mana keberhasilan guru/dosen sebagai fasilitator dalam menyajikan bahan pelajaran serta sejauh mana peserta didik dapat menyerap materi yang diajarkan.

Objektivitas penilaian hasil belajar yang dilakukan sangat bergantung pada kualitas dari alat ukur atau perangkat tes yang digunakan [8]. Umar (dalam Sehar [8]) menyatakan bahwa, hanya penilaian yang didasarkan kepada hasil pengukuran yang dapat dipercaya sajalah yang

¹Jurusan Matematika FMIPA Universitas Hasanuddin, email: nkalondeng@gmail.com

dapat dijadikan sebagai landasan yang kuat bagi pengambilan keputusan tentang keberhasilan pengajar maupun peserta didik. Namun dalam penerapannya, sering terjadi dua atau beberapa perangkat tes sering diperlakukan sama, tanpa memperhatikan tingkat kesukaran perangkat tes yang digunakan. Suryabrata [9] berpendapat bahwa dalam pelaksanaan evaluasi yang menggunakan beberapa perangkat tes yang berbeda namun mengukur hal yang sama perlu dilakukan penyetaraan skor dari beberapa perangkat tes tersebut, karena dengan adanya penyetaraan prinsip keadilan dalam ujian akan diperoleh peserta tes. Sebagaimana Hamblethton & Swaminathan (dalam [8]) mengingatkan bahwa sekalipun dua perangkat tes disusun berdasarkan kisi-kisi yang sama, jarang sekali bahkan hampir tidak ditemukan dua perangkat tes yang benar-benar setara dalam sebaran serta peringkat indeks kesukaran. Dengan demikian, menjadi suatu keharusan bagi para pembuat perangkat tes atau lembaga pengembang tes untuk melakukan penyetaraan terhadap perangkat yang digunakan.

Rahmat dalam [7] menyatakan bahwa untuk melihat kualitas sebuah perangkat tes dapat dilakukan dengan analisis butir per butir soal, atau yang lazim disebut dengan *item*. Tujuan dilakukannya analisis item antara lain untuk mengkaji dan menelaah setiap butir soal (*item*) agar diperoleh soal yang bermutu sebelum digunakan. Tujuan lain analisis item adalah membantu meningkatkan kualitas tes melalui revisi atau membuang soal yang tidak efektif, dan mengetahui informasi diagnostik pada siswa, sudahkah mereka memahami materi yang telah diajarkan atau belum. Sedangkan Misbah [6] menyatakan bahwa tujuan dari analisis *item* adalah untuk mengetahui apakah butir-butir soal yang membangun tes hasil belajar itu sudah dapat menjalankan fungsinya sebagai alat pengukur hasil belajar yang memadai ataukah belum, dimana parameter butir sebagai indikator yang digunakan adalah tingkat kesukaran, daya pembeda dan terkaan semu. Sehar [8] menyatakan bahwa indikator tingkat kesukaran dan daya pembeda umumnya digunakan pada soal isian atau *essay*, sedangkan terkaan semu umumnya digunakan untuk soal pilihan ganda.

Ada dua pendekatan yang dikembangkan untuk penilaian hasil belajar peserta didik melalui perangkat tes yang digunakan, yaitu pendekatan klasik hingga modern. Pendekatan secara klasik yang digunakan adalah teori tes klasik atau dikenal dengan *classical true-score theory*, sedangkan pendekatan modern dengan teori respon butir atau *item response theory* [10]. Teori tes klasik digunakan karena mudah dalam penerapannya, namun memiliki keterbatasan dalam mengukur tingkat kesukaran dan daya beda *item* karena perhitungan dua indikator tersebut didasarkan langsung pada skor total peserta tes. Sedangkan pendekatan modern dengan teori respon butir membebaskan ketergantungan antara *item* tes dan peserta tes (konsep invariansi parameter), respon peserta tes pada satu *item* tes tidak mempengaruhi *item* tes lainnya (konsep independensi lokal), dan *item* tes hanya mengukur satu dimensi ukur (konsep unidimensional). Sehingga aplikasinya menjawab kebutuhan dunia pengukuran modern hingga saat ini, yaitu perbandingan antar kemampuan peserta, pengembangan bank *item*, bahkan pengembangan tes adaptif, sehingga dianggap mampu mengatasi keterbatasan yang ada pada teori tes klasik (dalam Vitaria [10]).

Penelitian ini merupakan penelitian awal dengan tujuan jangka panjang untuk mengembangkan bank soal matakuliah Matematika Dasar yang berkualitas untuk mahasiswa baru di Universitas Hasanuddin. Sebagai tahap awal, penelitian ini difokuskan pada mahasiswa Fakultas Kehutanan Universitas Hasanuddin, dan mahasiswa Fakultas Ilmu Kelautan dan Perikanan (FIKP) yang mengambil matakuliah Matematika Dasar tahun ajaran 2011/2012, yang kebetulan diampu oleh peneliti. Penelitian ini juga merupakan salah satu upaya untuk mengembangkan kreativitas dengan mencoba menerapkan ilmu statistika dalam menyelesaikan permasalahan di dunia nyata.

2. Teori Respon Butir

Teori respon butir atau *Item Response Theory* (IRT), sehingga selanjutnya akan disingkat dengan IRT, merupakan teori pengukuran modern yang digunakan dalam menganalisis *item*. Teori ini mempunyai orientasi pada *item* yang karakteristiknya tidak tergantung pada kelompok tertentu. IRT membebaskan ketergantungan antara *item* tes dan peserta tes (konsep invariansi parameter), respon peserta tes pada satu *item* tes tidak mempengaruhi *item* tes lainnya (konsep independensi lokal), dan *item* tes hanya mengukur satu dimensi ukur (konsep unidimensional). Sehingga aplikasinya menjawab kebutuhan dunia pengukuran modern hingga saat ini, yaitu perbandingan antar kemampuan peserta dan pengembangan bank *item*, bahkan pengembangan tes adaptif, dan lain-lain [10]. Hambleton *et al.* dalam [4] mengemukakan bahwa asumsi-asumsi yang melandasi teori respons butir adalah:

1. Unidimensi, setiap butir hanya mengukur satu ciri peserta.
2. Independensi lokal, respon pada butir (*item*) yang satu bebas dari pengaruh respon pada butir lain jika kemampuan yang mempengaruhi performansi dibuat konstan.
3. Fungsi karakteristik butir atau kurva karakteristik butir, merefleksikan hubungan yang sebenarnya antara kemampuan dan respon peserta terhadap butir tes.

Kelemahan teori respons butir terletak pada tiga hal yaitu pemahaman, penghitungan, dan asumsi yang harus dipenuhi. Menurut Mardapi (dalam [5]), analisis butir tidak bisa dilakukan dengan manual tetapi harus menggunakan paket program komputer karena kompleksnya perhitungan. Setiap model menggunakan asumsi yang berbeda walaupun di dalamnya ada asumsi yang sama. Semakin banyak parameter dalam suatu model semakin kecil asumsi yang harus dipenuhi namun penghitungannya semakin kompleks.

2.1 Penerapan Teori Respons Butir dalam Penyetaraan Tes

Dalam penerapannya pada kegiatan penyetaraan tes asumsi-asumsi yang melandasi teori respon butir harus terpenuhi antara lain unidimensi dan independensi lokal [4]. Unidimensi artinya bahwa dimensi karakter peserta yang diukur oleh suatu tes itu tunggal. Independensi lokal adalah bahwa apabila kemampuan-kemampuan yang mempengaruhi kinerja tes dianggap konstan maka respons subjek terhadap setiap butir secara statistik tidak saling terkait.

Berbagai rancangan penyetaraan tes yang dapat digunakan menurut teori respons butir adalah:

1. Rancangan kelompok tunggal (*single-group design*)
Kegiatan penyetaraan dilakukan dengan menggunakan satu kelompok peserta yang merespons dua perangkat tes misalnya X dan Y. Parameter butir dari kedua perangkat tes diestimasi secara terpisah dengan mengkalibrasi parameter kemampuan peserta atau parameter butir.
2. Rancangan kelompok ekuivalen (*equivalent-group design*)
Desain ini merupakan kebalikan dari desain pertama, yaitu dua perangkat tes diberikan pada dua kelompok (X dan Y) dengan asumsi memiliki kemampuan yang sama atau ekuivalen. Proses secara spiral digunakan dalam desain ini, dimana peserta tes dibagi dua secara acak kemudian masing-masing mendapat perangkat tes 1 dan perangkat tes 2. Keuntungan dari rancangan ini adalah dapat menghindari efek negatif yang disebabkan karena adanya latihan dan kelelahan peserta tes, sedangkan kekurangannya adalah adanya kemungkinan bias yang disebabkan oleh perbedaan distribusi kemampuan dari kedua kelompok peserta tes.
3. Rancangan tes jangkar (*anchor test design*).

Pada desain ini masing-masing perangkat tes mempunyai beberapa butir soal yang sama (*common item*) dan masing-masing kelompok mengerjakan perangkat tes yang berbeda. Pada desain ini terdapat dua variasi yakni pertama, jika *common item* diperhitungkan dalam pemberian skor disebut *internal common item* dan kedua, jika *common item* tidak diperhitungkan dalam pemberian skor disebut *external common item*.

Berdasarkan rancangan di atas, maka penelitian ini akan difokuskan pada penggunaan rancangan kedua, dengan pertimbangan bahwa rancangan inilah yang sesuai dengan kondisi perkuliahan matakuliah Matematika Dasar di universitas ini.

2.2 Parameter Butir dalam Teori Respon Butir (IRT)

Parameter butir dalam Teori respon Butir (*Item response Theory*) yaitu, tingkat kesukaran *item* (b), daya beda *item* (a), peluang tebakan semu (c). Namun demikian, karena bentuk soal untuk Matematika Dasar umumnya dalam bentuk uraian atau isian, maka penelitian ini akan difokuskan pada parameter indeks daya beda butir/item (a) dan tingkat kesukaran butir/item (b).

Masing-masing parameter butir tersebut dapat dihitung dengan menggunakan rumus:

$$a_i = \frac{[(BA)_i - (BB)_i]}{\frac{N}{2}} \quad (1)$$

$$b_i = \frac{[(BA)_i + (BB)_i]}{N} \quad (2)$$

dimana,

a_i = Daya beda butir ke- i

b_i = Tingkat kesukaran butir ke- i

$(BA)_i$ = Jumlah benar kelompok atas butir ke- i

$(BB)_i$ = Jumlah benar kelompok bawah butir ke- i

N = Jumlah kelompok atas dan kelompok bawah

Penentuan kelompok atas dan bawah diambil 27 % dari jumlah responden yang telah diurutkan skornya dari skor tertinggi ke skor terendah [1].

Menurut Witherington dalam [8], angka indeks kesukaran butir (b) itu besarnya berkisar antara 0,00 sampai dengan 1,00. untuk menentukan tingkat kesukaran soal menurut Suharsimi (dalam [8]), ketentuan yang dilakukan adalah sebagai berikut:

- a. Soal dengan indeks kesukaran 0,00 sampai 0,30 adalah soal sukar
- b. Soal dengan indeks kesukaran 0,31 sampai 0,70 adalah soal sedang
- c. Soal dengan indeks kesukaran 0,71 sampai 1,00 adalah soal mudah.

Seperti halnya indeks kesukaran soal, maka indeks daya pembeda soal besarnya berkisar antara 0,00 sampai dengan 1,00 (Sudijono dalam [8]). Kriteria daya pembeda yakni:

- a. Soal dengan indeks daya pembeda 0,70 sampai 1,00 adalah baik sekali
- b. Soal dengan indeks daya pembeda 0,40 sampai 0,69 adalah baik
- c. Soal dengan indeks daya pembeda 0,20 sampai 0,39 adalah cukup
- d. Soal dengan indeks daya pembeda 0,00 sampai 0,19 adalah jelek
- e. Soal dengan indeks daya pembeda negatif (minus) adalah jelek sekali

3. Metode Penelitian

3.1 Jenis dan Sumber Data

Adapun jenis data yang digunakan dalam penelitian ini adalah data primer berupa hasil MID tes mahasiswa Fakultas Kehutanan dan FIKP Unhas menggunakan dua perangkat tes yang berbeda. Pengambilan Data direncanakan pada minggu pertama dan kedua September 2012.

3.2 Populasi/Sampel

Populasi penelitian ini adalah nilai MID tes semua mahasiswa Fakultas Kehutanan yang memprogramkan mata kuliah Matematika Dasar pada Semester Akhir Tahun Ajaran 2010/2011 dan mahasiswa FIKP yang memprogramkan mata kuliah Matematika Dasar pada Semester Akhir Tahun Ajaran 2011/2012 ini. Sampel yang terpilih adalah semua mahasiswa yang berada pada golongan tertentu yang berdasarkan pengaturan yang telah dilakukan oleh pengelola MKU, yang telah terpilih secara acak untuk menjadi fokus dalam penelitian ini. Pengambilan sampel dilakukan secara purposif yang disesuaikan dengan tujuan penelitian.

3.3 Indikator/Parameter yang Diukur

Indikator/parameter yang diamati dalam penelitian ini diberikan pada tabel berikut.

Tabel 1. Indikator/Parameter yang diukur.

Pengamatan	Jenis Variabel/Peubah
Skor Tiap Soal untuk Nilai Mata Kuliah Matematika Dasar untuk Kelompok K1 (X)	Kontinu
Skor Tiap Soal untuk Nilai Mata Kuliah Matematika Dasar untuk Kelompok K2 (Y)	Kontinu
Skor Tiap Soal untuk Nilai Mata Kuliah Matematika Dasar untuk Kelompok L1 (X)	Kontinu
Skor Tiap Soal untuk Nilai Mata Kuliah Matematika Dasar untuk Kelompok L2 (Y)	Kontinu

3.4 Metode Analisis Data

Adapun langkah-langkah yang akan dilakukan, yang berkaitan dengan tujuan penelitian ini adalah sebagai berikut:

1. Menentukan rancangan penarikan sampel yaitu rancangan ekuivalen (*equivalent-group design*).
2. Menghitung indeks parameter butir yaitu tingkat kesukaran (*b*) dan daya beda (*a*).
3. Menganalisis butir-butir soal perangkat tes X dan Y berdasarkan indeks parameter butir.
4. Melakukan uji homogenitas varians untuk mengetahui apakah kedua perangkat tes sudah setara atau belum.

4. Hasil dan Pembahasan

Beberapa hasil telah diperoleh dari penelitian ini, namun sebelum mengemukakan hasil tersebut akan dijelaskan beberapa hal mendasar yang terkait dengan analisis respon terhadap butir soal sampai diperolehnya hasil penelitian ini.

Sebagaimana yang telah dikemukakan pada bagian sebelumnya, bahwa rancangan penarikan sampel untuk penelitian ini adalah menggunakan rancangan ekivalen, dimana sebelum dilakukan ujian MID tes, mahasiswa fakultas Kehutanan yang memprogramkan mata kuliah Matematika Dasar pada Semester Akhir Tahun Ajaran 2010/2011 dibagi kedalam dua kelompok yang ekivalen, misalkan dinamakan dengan K1 dan K2. Kelompok K1 yang mengerjakan tes X terdiri dari 32 orang mahasiswa, dan K2 yang mengerjakan tes Y juga terdiri dari 32 orang mahasiswa. Hal yang sama juga dilakukan terhadap mahasiswa FIKP yang memprogramkan mata kuliah Matematika Dasar pada Semester Akhir Tahun Ajaran 2011/2012, juga dibagi ke dalam 2 kelompok yang ekivalen, misalnya dinamakan kelompok L1 dan L2, dimana kelompok L1 yang mengerjakan tes X terdiri dari 26 orang mahasiswa, dan L2 yang mengerjakan tes Y terdiri dari 26 orang mahasiswa. Total jumlah butir soal pada masing-masing perangkat tes adalah 13 butir soal uraian atau esai, dengan skor tertinggi tiap butir adalah 10 dan skor terendah adalah 0. Penelitian pendahuluan yang dilakukan oleh peneliti dengan beberapa mahasiswa bimbingan juga telah diperoleh hasil bahwa kedua perangkat tes yang digunakan, yaitu perangkat tes X dan Y telah diuji kesetaraannya dan disimpulkan bahwa kedua perangkat tersebut adalah setara.

Dalam menghitung indeks parameter butir daya beda soal (a) dan tingkat kesukaran (b), data skor peserta yang diperoleh pada penelitian dikonversi sesuai dengan banyaknya kategori yang ditetapkan dalam penilaian suatu butir soal, sehingga nantinya metode analisis butir yang digunakan disesuaikan dengan banyaknya kategori penilaian yang telah dipilih. Dalam penelitian ini, akan dicobakan dua bentuk konversi nilai tiap butir, dimana konversi pertama digunakan konversi nilai kedalam bilangan biner 0 dan 1, dengan ketentuan jika peserta memperoleh nilai < 5 diberi kode 0, dan sebaliknya jika peserta memperoleh nilai ≥ 5 diberi kode 1, sehingga model/metode analisis butir yang akan digunakan adalah metode analisis yang paling umum untuk konversi ke dalam dua kategori (biner) dengan menggunakan model Regresi. Penskoran dengan cara seperti ini disebut dengan Penskoran Dikotomi. Metode ini terutama akan digunakan untuk menganalisis skor nilai mid tes tiap butir soal mahasiswa Fakultas Kehutanan. Sedangkan untuk mahasiswa FIKP digunakan konversi kedua dengan Penskoran Politomi, dimana skor nilai peserta dibagi ke dalam 3 kategori dengan ketentuan peserta yang memperoleh skor 0-4 akan dikonversi ke 0, peserta yang memperoleh skor 5-7 akan dikonversi ke 1, dan peserta yang memperoleh skor 8-10 akan dikonversi ke 2. Model yang akan digunakan untuk Penskoran Politomi ini dipilih *Generalized Partial Credit Model* (GPCM) yang merupakan perluasan dari *Partial Credit Model* (PCM) untuk beberapa indeks parameter butir. Sehingga hasil dan pembahasan dari penelitian ini akan dibagi ke dalam dua bagian berikut.

4.1. Penskoran Dikotomi dengan Model Regresi.

Hasil penelitian dengan menggunakan beberapa langkah yang dilakukan dalam Penskoran Dikotomi dengan model Regresi akan diuraikan berikut.

1. Penentuan Jumlah Benar Kelompok Atas (BA) dan Kelompok Bawah(BB)

Seperti telah dikemukakan sebelumnya bahwa perangkat tes yang digunakan terdiri dari perangkat tes X yang dikerjakan oleh kelompok K1 dan perangkat tes Y yang dikerjakan oleh K2. Pada masing-masing kelompok kemudian ditentukan jumlah peserta menjawab benar pada

kelompok atas (BA) dan jumlah peserta menjawab benar pada kelompok bawah (BB) setelah diurutkan skor peserta dari yang terbesar ke yang terkecil. Penentuan skor peserta tes yang masuk dalam kelompok atas (BA) dan kelompok bawah (BB) dihitung dengan menggunakan rumus:

$$\frac{27}{100} \times 32 = 8,64 \text{ dibulatkan menjadi } 9$$

sehingga nilai skor peserta yang berada pada nomor urut 1-9 setelah skor awal diurutkan dari yang terbesar ke yang terkecil, dimasukkan kedalam kelompok atas (BA), dan 9 urutan terakhir skor peserta dimasukkan dalam kelompok bawah (BB), baik pada K1 maupun pada K2, sebagaimana yang ditunjukkan pada Tabel 2a, 2b, 2c, dan 2d berikut.

Tabel 2a menunjukkan skor butir peserta dalam kelompok K1 yang masuk dalam BA setelah peserta dites menggunakan perangkat tes X, sedangkan Tabel 2b skor butir peserta dalam kelompok K1 yang masuk dalam BA setelah peserta dites menggunakan perangkat tes Y. Tabel 2c menunjukkan skor butir peserta dalam kelompok K1 yang masuk dalam BB setelah peserta dites menggunakan perangkat tes X, sedangkan Tabel 2d skor butir peserta dalam kelompok K1 yang masuk dalam BB setelah peserta dites menggunakan perangkat tes Y.

Tabel 2a. Skor Tes Butir Soal Peserta dalam Kelompok Atas K1 dengan Perangkat Tes X.

Peserta pada Kelompok Atas K1	Butir Soal													Skor
	1	2	3	4	5	6	7	8	9	10	11	12	13	
1	1	0	1	1	1	0	0	0	0	1	1	0	0	6
2	1	0	1	1	0	0	0	0	0	1	1	0	1	6
3	1	1	1	1	0	0	0	0	0	1	0	0	0	5
4	1	0	1	1	0	0	1	0	0	0	1	0	0	5
5	1	1	1	1	0	0	0	0	0	0	0	0	0	4
6	1	1	1	1	0	0	0	0	0	0	0	0	0	4
7	0	0	1	1	0	0	1	0	0	0	1	0	0	4
8	1	0	0	1	0	0	1	0	0	0	1	0	0	4
9	1	0	1	1	0	0	0	0	0	0	0	0	0	3
Jumlah	8	3	8	9	1	0	3	0	0	3	5	0	1	

Tabel 2b. Skor Tes Butir Soal Peserta dalam Kelompok Atas K2 dengan Perangkat Tes Y.

Peserta Pada Kelompok Atas K2	Butir Soal													Skor
	1	2	3	4	5	6	7	8	9	10	11	12	13	
1	1	1	0	1	1	0	1	0	1	0	0	0	1	7
2	1	0	0	1	1	0	1	0	0	0	1	0	1	6
3	1	0	0	1	1	0	1	0	0	0	1	0	0	5
4	1	1	0	0	1	0	1	0	0	0	1	0	1	6
5	1	1	0	0	1	0	1	0	0	0	1	0	1	6
6	1	0	0	1	1	0	1	0	0	0	1	0	0	5
7	1	0	0	1	0	0	1	0	0	0	1	0	0	4
8	1	0	0	1	0	0	1	0	0	0	1	0	0	4
9	1	0	1	1	1	0	1	0	0	0	0	0	1	6
Jumlah	9	3	1	7	7	0	9	0	1	0	7	0	5	

Tabel 2c. Skor Tes Butir Soal Peserta dalam Kelompok Bawah K1 dengan Perangkat Tes X.

Peserta pada Kelompok Bawah K1	Butir Soal													Skor
	1	2	3	4	5	6	7	8	9	10	11	12	13	
1	1	0	0	0	0	0	1	0	0	1	0	0	0	3
2	1	0	0	0	0	0	1	0	0	0	1	0	0	3
3	1	1	0	0	0	0	0	0	0	0	0	0	0	2
4	1	0	0	0	0	0	1	0	0	1	0	0	0	3
5	1	0	0	0	0	0	1	0	0	0	0	0	0	2
6	1	0	0	0	0	0	1	0	0	0	0	0	0	2
7	0	0	0	1	0	0	0	0	0	0	0	0	0	1
8	1	0	0	0	0	0	0	0	0	0	0	0	0	1
9	1	0	0	0	0	0	0	0	0	0	0	0	0	1
Jumlah	8	1	0	1	0	0	5	0	0	2	1	0	0	

Tabel 2d. Skor Tes Butir Soal Peserta dalam Kelompok Bawah K2 dengan Perangkat Tes Y.

Peserta pada Kelompok Bawah K2	Butir Soal													Skor
	1	2	3	4	5	6	7	8	9	10	11	12	13	
1	1	1	0	1	0	0	1	0	0	0	0	0	0	4
2	1	0	0	1	0	0	1	0	0	0	0	0	0	3
3	1	0	0	1	0	0	0	0	0	0	0	0	0	2
4	1	0	0	1	0	0	1	0	0	0	0	0	0	3
5	1	1	0	0	0	0	0	0	0	0	0	0	0	2
6	0	0	0	1	0	0	0	0	0	0	0	0	0	1
7	0	0	0	1	0	0	0	0	0	0	0	0	0	1
8	1	0	0	0	1	0	0	0	0	0	0	0	0	2
9	0	0	0	1	0	0	0	0	0	0	0	0	0	1
Jumlah	6	2	0	7	1	0	3	0	0	0	0	0	0	

Setelah penentuan nilai jumlah benar kelompok atas (BA) dan jumlah benar kelompok bawah (BB), langkah selanjutnya adalah analisis butir untuk menentukan nilai daya beda (a) dan tingkat kesukaran (b).

2. Perhitungan Indeks Daya Beda

Langkah selanjutnya adalah menghitung indeks daya beda (a) kedua perangkat tes yang digunakan, yaitu perangkat tes X dan Y, dimana dalam perhitungannya menggunakan data yang diberikan pada Tabel 2a, 2b, 2c, dan 2d. Hasil perhitungan indeks daya beda untuk setiap butir soal diberikan pada Tabel 3 berikut. Pada tabel yang sama ditunjukkan juga kategori setiap butir soal, apakah termasuk soal yang baik sekali, baik, cukup, jelek atau jelek sekali dalam membedakan kemampuan peserta tes sehingga menjadi rekomendasi sebagai soal yang perlu direvisi atau ditinjau kembali, bahkan jika perlu dibuang.

Nilai daya beda perangkat tes X dan perangkat tes Y pada Tabel 3 dihitung dengan menggunakan rumus pada persamaan (1). Contoh perhitungan secara manual daya beda soal Nomor 1 pada perangkat tes X jika diketahui $(BA)_i = 8$, $(BB)_i = 8$ dan $N = 18$; maka

$$a_i = \frac{[(BA)_i - (BB)_i]}{\frac{N}{2}}$$

$$a_i = \frac{(8 - 8)}{\frac{18}{2}}$$

$$a_i = 0$$

Tabel 3. Nilai Daya Beda Perangkat Tes X dan Perangkat Tes Y dan Kategori Masing-Masing Soal.

Butir Soal	Daya Beda (a)Perangkat Tes X	Kategori Butir Soal	Daya Beda (a)Perangkat Tes Y	Kategori Butir Soal
1	0	Jelek	0.33	Cukup
2	0.22	Jelek	0.11	Jelek
3	0.89	Baik Sekali	0.11	Jelek
4	0.89	Baik Sekali	0	Jelek
5	0.11	Jelek	0.67	Baik
6	0	Jelek	0	Jelek
7	-0.22	Jelek Sekali	0.67	Baik
8	0	Jelek	0	Jelek
9	0	Jelek	0.11	Jelek
10	0.11	Jelek	0	Jelek
11	0.44	Baik	0.78	Baik Sekali
12	0	Jelek	0	Jelek
13	0.11	Jelek	0.56	Baik

Jadi nilai daya beda soal nomor 1 untuk kelompok K1 yang menggunakan perangkat tes X adalah 0, sehingga soal Nomor 1 dikategorikan sebagai soal yang jelek. Soal ini tidak mampu membedakan antara mahasiswa yang pintar, sedang ataupun yang kurang. Sehingga saran perbaikan atau melihat kembali butir soal ini perlu dilakukan. Soal pada perangkat tes X yang perlu mendapat perhatian kembali untuk bisa digunakan kembali sebagai alat tes kemampuan peserta karena daya bedanya tidak begitu baik dalam membedakan kemampuan mahasiswa adalah soal nomor 2, 5, 6, 7, 8, 9, 10, 12, dan 13. Sedangkan untuk butir soal pada perangkat tes Y yang perlu untuk ditinjau kembali karena mempunyai daya beda kecil dalam membedakan kemampuan mahasiswa adalah soal Nomor 2, 3, 4, 6, 8, 9, 10, dan 12.

3. Perhitungan Indeks Tingkat Kesukaran Butir Soal

Indeks tingkat kesukaran butir mengukur respon peserta tes terhadap suatu butir soal, apakah butir tersebut termasuk ke dalam soal yang sulit, sedang atau mudah. Nilai indeks butir yang kecil menunjukkan bahwa butir soal tersebut tergolong sulit untuk dikerjakan oleh peserta tes, sedangkan nilai indeks yang besar menunjukkan bahwa butir soal tersebut tergolong mudah untuk dikerjakan oleh semua peserta tes. Hasil perhitungan nilai indeks tingkat kesukaran untuk kedua kelompok K1 dan K2 serta perangkat tes X dan Y yang digunakan, diberikan pada Tabel 4

berikut. Pada tabel tersebut juga diberikan kategori tingkat kesukaran setiap butir pada kedua perangkat tes.

Contoh perhitungan manual tingkat kesukaran butir (b) pada soal Nomor 1 perangkat tes X yang diberikan kepada kelompok K1 adalah sebagai berikut.

Jika diketahui $(BA)_i = 8$, $(BB)_i = 8$ dan $N = 18$; maka

$$b_i = \frac{[(BA)_i + (BB)_i]}{N}$$

$$b_i = \frac{(8+8)}{18}$$

$$b_i = 0,83$$

Tabel 4. Nilai Indeks Tingkat Kesukaran Butir Perangkat Tes X dan Perangkat Tes Y dan Kategori Masing-Masing Soal.

Butir Soal	Tingkat Kesukaran Butir (b) Perangkat Tes X	Kategori Butir Soal	Tingkat Kesukaran Butir (b) Perangkat Tes Y	
			Kategori Butir Soal	Kategori Butir Soal
1	0.89	Mudah	0.83	Mudah
2	0.22	Sulit	0.28	Sulit
3	0.44	Sedang	0.06	Sulit
4	0.56	Sedang	0.78	Mudah
5	0.06	Sulit	0.44	Sedang
6	0	Sulit	0	Sulit
7	0.44	Sedang	0.67	Sedang
8	0	Sulit	0	Sulit
9	0	Sulit	0.06	Sulit
10	0.28	Sulit	0	Sulit
11	0.33	Sedang	0.39	Sedang
12	0	Sulit	0	Sulit
13	0.06	Sulit	0.28	Sulit

Jadi nilai tingkat kesukaran soal Nomor 1 dengan perangkat tes X adalah 0.89. Hal ini berarti bahwa soal Nomor 1 termasuk kategori mudah dikerjakan oleh setiap peserta, sehingga soal ini bisa direkomendasikan untuk bisa digunakan dalam tes untuk peserta yang memprogramkan matakuliah ini selanjutnya. Soal-soal yang perlu mendapat perhatian kembali atau revisi lanjut adalah menurut indeks tingkat kesukaran adalah soal Nomor 2, 5, 6, 8, 9, 10, 12, dan 13 pada perangkat tes X dan soal Nomor 2, 3, 6, 8, 9, 10, 12, dan 13 pada perangkat tes Y.

4. Estimasi Persamaan Penyetaraan dengan Model Regresi

Dalam bagian ini akan dilakukan konversi dari indeks daya beda dan tingkat kesukaran butir pada perangkat tes X ke perangkat tes Y. hal ini dimaksudkan untuk mengetahui bagaimana indeks daya beda dan tingkat kesukaran setiap butir soal peserta tes pada kelompok K1 seandainya diberikan perangkat tes Y. Hal sebaliknya bisa dilakukan pada peserta tes yang berada di kelompok K2.

Dengan menggunakan indeks daya beda (a) dan tingkat kesukaran butir (b) pada perangkat tes X dan Y maka konversi nilai daya beda X (aX) ke nilai daya beda Y ($a*Y$) dan nilai tingkat kesukaran X (bX) ke nilai tingkat kesukaran Y ($b*Y$) dapat dilakukan, dan hasilnya diberikan pada Tabel 5 berikut. Nilai $a * Y$ diperoleh melalui penentuan konstanta konversi α dan β menggunakan metode regresi yang dilakukan dengan memperhatikan respons peserta tes pada kedua perangkat tes X dan Y. Estimasi parameter butir memenuhi persamaan model regresi linier sebagai berikut:

$$y = \alpha x + \beta + \varepsilon \quad (3)$$

dengan model penduga

$$\hat{y} = \hat{\alpha}x + \hat{\beta} \quad (4)$$

$$\hat{\alpha} = \frac{r_{xy}s_y}{s_x} \quad (5)$$

$$\hat{\beta} = \bar{y} - \alpha\bar{x} \quad (6)$$

dimana:

\hat{y} = Estimasi kemampuan atau estimasi parameter butir pada perangkat tes Y,

\hat{x} = Estimasi kemampuan atau estimasi parameter butir pada perangkat tes X,

\bar{y}, \bar{x} = Rerata dari y dan x,

r_{XY} = Koefisien korelasi antara X dan Y,

$s_x s_y$ = Simpangan baku dari x dan y.

Hasil perhitungan konversi untuk indeks daya beda dan tingkat kesukaran butir diberikan pada Tabel 5 berikut.

Berdasarkan indeks daya beda, dari tabel di atas terlihat bahwa beberapa beberapa soal sangat tergantung pada kemampuan peserta tes, meskipun dari awal pembentukan kelompok K1 dan K2 diasumsikan bahwa kedua kelompok peserta tes tersebut adalah setara, artinya kemampuan peserta tes di kelompok K1 relatif sama dengan kemampuan peserta tes di kelompok K2. Asumsi ini dijamin oleh pembentukan kelompok yang mempertimbangkan faktor kesetaraan dalam kemampuan peserta tes ini. Namun respon peserta tes menunjukkan bahwa ada perbedaan kemampuan peserta tes dalam merespon suatu butir soal, seperti yang terlihat pada soal butir 1 misalnya. Peserta tes di kelompok K1 memberikan respon bahwa soal nomor ini jelek dalam membedakan kemampuan peserta, sementara konversi dari peserta dalam kelompok K2 memperlihatkan bahwa soal Nomor 1 ini cukup dalam membedakan kemampuan peserta tes. Hal yang sama terlihat pada butir lain seperti Nomor 2, 3, 4, 7, 8, 9, 10, 11, 12, dan 13. Sedangkan soal Nomor 5 dan 6 direkomendasikan untuk direvisi atau jika perlu dibuang karena butir soal tersebut jelek dalam membedakan kemampuan peserta baik di kelompok K1 maupun K2.

Tabel 5. Hasil Konversi Daya Beda aX ke $a*Y$ dan Tingkat Kesukaran bX ke $b*Y$.

Butir Soal X	aX	$a*Y$	Kategori Konversi	bX	$b*Y$	Kategori Konversi
1	0	0.29	Jelek → Cukup	0.89	0.83	Mudah → Mudah
2	0.22	0.25	Cukup → Cukup	0.22	0.26	Sulit → Sulit
3	0.89	0.16	Baik Sekali → Jelek	0.44	0.45	Sedang → Sedang
4	0.89	0.16	Baik Sekali → Jelek	0.56	0.55	Sedang → Sedang
5	0.11	0.27	Jelek → Jelek	0.06	0.13	Sulit → Sulit
6	0	0.29	Jelek → Jelek	0	0.08	Sulit → Sulit
7	-0.22	0.32	Jelek Sekali → Cukup	0.44	0.45	Sedang → Sedang
8	0	0.29	Jelek → Cukup	0	0.08	Sulit → Sulit
9	0	0.29	Jelek → Cukup	0	0.08	Sulit → Sulit
10	0.11	0.27	Jelek → Cukup	0.28	0.32	Sulit → Sedang
11	0.44	0.22	Baik → Cukup	0.33	0.36	Sedang → Sedang
12	0	0.29	Jelek → Cukup	0	0.08	Sulit → Sulit
13	0.11	0.27	Jelek → Cukup	0.06	0.13	Sulit → Sulit

Berdasarkan indeks tingkat kesukaran butir, tabel di atas memperlihatkan tidak ada perbedaan respon antara peserta pada kelompok K1 dan K2. Artinya soal yang mudah menurut peserta pada kelompok K1 maka peserta K2 akan merespon seperti peserta pada kelompok K1 juga.

4.2 Penskoran Politomi dengan *Generalized Partial Credit Model* (GPCM).

Untuk menentukan parameter butir, data skor mahasiswa FIKP yang diperoleh pada penelitian ini dikonversi kedalam bentuk penskoran politomi 0,1 dan 2. Peserta yang memperoleh skor 0-4 akan dikonversi ke 0, peserta yang memperoleh skor 5-7 akan dikonversi ke 1, dan peserta yang memperoleh skor 8-10 akan dikonversi ke skor 2.

Seperti yang telah dikemukakan bahwasanya perangkat tes yang digunakan terdiri dari perangkat tes X yang dikerjakan oleh kelompok K1 dan perangkat tes Y yang dikerjakan oleh kelompok K2x. Pada masing-masing kelompok kemudian ditentukan jumlah peserta yang menjawab benar pada kelompok atas (A) dan jumlah peserta yang menjawab benar pada kelompok bawah (B). Untuk menentukan peserta kelompok atas (A) dan peserta kelompok bawah diperoleh dengan cara mengurutkan skor peserta ujian dari yang paling tinggi sampai yang paling rendah kemudian diambil 27% teratas dan terbawah dari jumlah peserta ujian untuk masing-masing kelompok (K_1) dan kelompok (K_2), yaitu $\frac{27}{100} \times 26 = 7,02$, dibulatkan menjadi 7 peserta.

1. Perhitungan Indeks Daya Beda

Untuk perhitungan indeks daya beda butir soal dengan menggunakan Penskoran Politomi, maka persamaan yang digunakan adalah

$$\alpha_i = \frac{A-B}{N(S_{Maks}-S_{Min})} \quad (7)$$

dimana :

- α_i = Daya pembeda soal pada butir soal ke
- A = Jumlah skor kelompok atas i
- B = Jumlah skor kelompok bawah
- N = Jumlah siswa kelompok atas atau bawah
- S_{Maks} = Skor tertinggi setiap soal uraian
- S_{Min} = Skor terendah setiap soal uraian

Hasil perhitungan indeks daya beda selengkapnya diberikan pada Tabel 6 berikut.

Jadi nilai daya beda soal nomor 1 untuk kelompok K1 yang menggunakan perangkat tes X adalah 1, sehingga soal Nomor 1 dikategorikan sebagai soal yang baik sekali. Soal ini mampu membedakan antara mahasiswa yang pintar, sedang ataupun yang kurang. Sehingga rekomendasi untuk soal ini adalah sebaiknya soal ini dimasukkan sebagai bank soal untuk jadi acuan pada tes yang akan dilakukan pada tahun-tahun selanjutnya. Soal pada perangkat tes X yang perlu mendapat perhatian untuk bisa digunakan kembali sebagai alat tes kemampuan peserta karena daya bedanya tidak begitu baik dalam membedakan kemampuan mahasiswa adalah soal nomor 3, 4, 6, 9, 10, 12, dan 13. Sedangkan soal lain bisa direkomendasikan untuk dimasukkan ke dalam bank soal, dengan melakukan revisi seperlunya jika dibutuhkan. Sedangkan untuk butir soal pada perangkat tes Y yang perlu untuk ditinjau kembali karena mempunyai daya beda kecil dalam membedakan kemampuan mahasiswa adalah soal Nomor 4, 7, dan 13, sedangkan yang lainnya direkomendasikan untuk dimasukkan sebagai bank soal.

2. Perhitungan Indeks Tingkat Kesukaran Butir Soal

Indeks tingkat kesukaran butir mengukur respon peserta tes terhadap suatu butir soal, apakah butir tersebut termasuk ke dalam soal yang sulit, sedang atau mudah. Pada bentuk butir politomi, nilai tingkat kesukaran butir adalah jumlah dari tingkat kesukaran tahap tiap butir, jadi berbeda cara perhitungannya dengan Penskoran Dikotomi. Persamaan yang digunakan untuk menghitung indeks ini adalah

$$\delta_{ij} = \frac{n_{ij}}{N_i} \quad (8)$$

dimana

- δ_{ij} = tingkat kesukaran tahap j pada butir i
- n_{ij} = jumlah skor yang diperoleh peserta yang menjawab pada butir i kategori j
- N_i = skor total pada butir i

Nilai indeks butir yang kecil menunjukkan bahwa butir soal tersebut tergolong sulit untuk dikerjakan oleh peserta tes, sedangkan nilai indeks yang besar menunjukkan bahwa butir soal tersebut tergolong mudah untuk dikerjakan oleh semua peserta tes. Hasil perhitungan nilai indeks tingkat kesukaran untuk kedua kelompok K1 dan K2 serta perangkat tes X dan Y yang digunakan, diberikan pada Tabel 7 berikut.

Tabel 6. Nilai Daya Beda Perangkat Tes X dan Perangkat Tes Y dan Kategori Masing-Masing Soal dengan GPCM.

Butir Soal	Daya Beda (a)Perangkat Tes X	Kategori Butir Soal	Daya Beda (a)Perangkat Tes Y	Kategori Butir Soal
1	1	Baik Sekali	0.5	Baik
2	0.21428571	Cukup	0.714285714	Baik Sekali
3	0.14285714	Jelek	0.428571429	Baik
4	0.14285714	Jelek	0.071428571	Jelek
5	0.28571429	Cukup	0.357142857	Cukup
6	0.14285714	Jelek	0.714285714	Baik Sekali
7	0.57142857	Baik	0.142857143	Jelek
8	0.64285714	Baik	0.642857143	Baik
9	0.14285714	Jelek	0.357142857	Cukup
10	0.14285714	Jelek	0.214285714	Cukup
11	0.42857143	Baik	0.571428571	Baik
12	0.14285714	Jelek	0.714285714	Baik Sekali
13	0	Jelek	0.142857143	Jelek

Contoh perhitungan nilai tingkat kesukaran butir soal untuk peserta pertama pada kelompok K1 yang mengerjakan perangkat tes X adalah sebagai berikut.

- 1) Tingkat kesukaran untuk soal nomor 1 kategori 0.

Diketahui :

$$n_{10} = 15 \times 0 = 0$$

$$N_1 = 52$$

Ditanyakan : $\delta_{10} = ?$

$$\delta_{10} = \frac{n_{10}}{N_1} = \frac{0}{52} = 0$$

- 2) Tingkat kesukaran untuk soal nomor 1 kategori 1.

Diketahui : $n_{11} = 0 \times 1 = 0$, $N_1 = 52$

Ditanyakan : $\delta_{11} = ?$

Penyelesaian :

Anisa

$$\delta_{11} = \frac{n_{11}}{N_1} = \frac{0}{52} = 0$$

- 3) Tingkat kesukaran untuk soal nomor 1 kategori 2.
 Diketahui : $n_{11} = 11 \times 2 = 22$, $N_1 = 52$
 Ditanyakan : $\delta_{12} = ?$
 Penyelesaian :

$$\delta_{11} = \frac{n_{12}}{N_1} = \frac{22}{52} = 0.423$$

Nilai tingkat kesukaran untuk butir soal nomor 1 kelompok K1 yang menggunakan perangkat tes X adalah:

$$P = \delta_{10} + \delta_{11} + \delta_{12} = 0 + 0 + 0.423 = 0.423$$

Untuk semua butir soal nilai $\delta_{10} = 0$. Dengan cara yang sama ditemukan nilai tingkat kesukaran pada nomor soal berikutnya, sebagaimana yang ditampilkan hasilnya pada Tabel 7 di bawah.

Tabel 7. Nilai Indeks Tingkat Kesukaran Butir Perangkat Tes X dan Perangkat Tes Y dan Kategori Masing-Masing Soal dengan GPCM.

Butir Soal	Tingkat Kesukaran Butir (b) Perangkat Tes X	Kategori Butir Soal	Tingkat Kesukaran Butir (b) Perangkat Tes Y	
			Kategori Butir Soal	Kategori Butir Soal
1	0.423	Sedang	0.173	Sulit
2	0.096	Sulit	0.404	Sedang
3	0.077	Sulit	0.192	Sulit
4	0.115	Sulit	0.25	Sulit
5	0.077	Sulit	0.212	Sulit
6	0.019	Sulit	0.25	Sulit
7	0.423	Sedang	0.654	Sedang
8	0.212	Sulit	0.327	Sedang
9	0.019	Sulit	0.135	Sulit
10	0.096	Sulit	0.077	Sulit
11	0.269	Sulit	0.269	Sulit
12	0.038	Sulit	0.481	Sedang
13	0	Sulit	0.019	Sulit

Dari Tabel 7 di atas terlihat bahwa beberapa butir soal sangat tergantung dari respon peserta terhadap soal tersebut, namun beberapa butir soal yang memperlihatkan bahwa butir soal tertentu

tergolong sulit pada kedua perangkat tes yang digunakan, yaitu perangkat tes X dan Y, sebaiknya ditinjau kembali dan dilakukan revisi seperlunya. Soal tersebut adalah soal butir 3, 4, 5, 6, 9, 10, 11 dan 13. Sedangkan soal yang lain berdasarkan tingkat kesukarannya dapat dimasukkan ke dalam bank soal untuk dapat digunakan pada tes-tes selanjutnya.

4.3 Perbandingan Hasil Penskoran Dikotomi dan Penskoran Politomi

Pada bagian ini akan dicoba untuk menguraikan hasil yang diperoleh dengan menggunakan Penskoran Dikotomi dan Penskoran Politomi pada dua gugus data penelitian yang digunakan dengan dua perangkat tes yaitu X dan Y, dengan memperhitungkan kemampuan peserta tes pada dua kelompok yang berbeda, untuk melihat butir soal mana saja yang bisa dipertimbangkan masuk dalam bank soal dan butir soal mana saja yang perlu direvisi dan lakukan perbaikan ataupun dibuang. Hasilnya diringkaskan dalam Tabel 8 berikut.

Berdasarkan hasil pada Tabel 8 direkomendasikan untuk memasukkan soal nomor 1 dan 7 ke dalam bank soal, soal nomor 2, 3, 4, 8, 9, 10, 11, dan 12 bisa dimasukkan ke dalam bank soal setelah terlebih dahulu dilakukan revisi terutama menyangkut tingkat kesulitan butir-butir soal tersebut, sedangkan soal nomor 5 dan 6 sebaiknya dibuang dan diganti dengan nomor lain.

Tabel 8. Perbandingan Hasil Penskoran Dikotomi dan Politomi terhadap Indeks Butir Soal yang Digunakan.

Butir Soal	Penskoran Dikotomi		Penskoran Politomi	
	Daya Beda (a)	Tingkat Kesukaran Butir (b)	Daya Beda (a)	Tingkat Kesukaran Butir (b)
1	Masuk bank soal dan lakukan revisi	Masuk bank soal	Masuk bank soal	Masuk bank soal dan lakukan revisi
2	Masuk bank soal dan lakukan revisi	Lakukan revisi	Masuk bank soal dan lakukan revisi	Masuk bank soal dan lakukan revisi
3	Masuk bank soal dan lakukan revisi	Masuk bank soal	Masuk bank soal dan lakukan revisi	Sebaiknya dibuang
4	Masuk bank soal dan lakukan revisi	Masuk bank soal	Sebaiknya dibuang	Sebaiknya dibuang
5	Sebaiknya dibuang	Sebaiknya dibuang	Masuk bank soal dan lakukan revisi	Sebaiknya dibuang
6	Sebaiknya dibuang	Sebaiknya dibuang	Masuk bank soal dan lakukan revisi	Sebaiknya dibuang
7	Masuk bank soal dan lakukan revisi	Masuk bank soal	Masuk bank soal dan lakukan revisi	Masuk bank soal
8	Masuk bank soal dan lakukan revisi	Lakukan revisi	Masuk bank soal	Masuk bank soal dan lakukan revisi

Butir Soal	Penskoran Dikotomi		Penskoran Politomi	
	Daya Beda (a)	Tingkat Kesukaran	Daya Beda (a)	Tingkat Kesukaran

		Butir (b)		Butir (b)
9	Masuk bank soal dan lakukan revisi	Lakukan revisi	Masuk bank soal dan lakukan revisi	Sebaiknya dibuang
10	Masuk bank soal dan lakukan revisi	Lakukan revisi	Masuk bank soal dan lakukan revisi	Sebaiknya dibuang
11	Masuk bank soal dan lakukan revisi	Lakukan revisi	Masuk bank soal	Sebaiknya dibuang
12	Masuk bank soal dan lakukan revisi	Lakukan revisi	Masuk bank soal dan lakukan revisi	Masuk bank soal dan lakukan revisi
13	Masuk bank soal dan lakukan revisi	Lakukan revisi	Sebaiknya dibuang	Sebaiknya dibuang

5. Kesimpulan

Penelitian ini merupakan penelitian awal mengenai metode analisis butir soal pada tes uraian untuk matakuliah Matematika Dasar. Hasil yang diperoleh adalah rekomendasi untuk menjadikan butir 1 dan 7 pada perangkat tes X dan Y yang digunakan sebagai bank soal, melihat kembali dan melakukan revisi pada butir soal 2, 3, 4, 8, 9, 10, 11, dan 12 pada kedua perangkat tersebut sebelum dimasukkan ke dalam bank soal, dan membuang dan mengganti soal nomor 5 dan 6.

Beberapa saran untuk penelitian lanjut ini bisa dilakukan baik dari segi pengembangan metode Penskoran Dikotomi dan Politomi yang digunakan, juga karakteristik peserta yang bisa jadi sangat mempengaruhi hasil dari penelitian ini. Sebagaimana yang diketahui, bahwa penelitian awal ini baru dicobakan pada mahasiswa dari Fakultas Kehutanan dan FIKP, yang merupakan salah satu bagian kecil dari peserta matakuliah Matematika Dasar di bawah Koordinasi dari UPT MKU. Penelitian lanjut juga dapat dilakukan dengan mempertimbangkan jumlah soal uraian yang lebih banyak dibanding soal yang digunakan sekarang.

Daftar Pustaka

- [1] Anonim. 2009. URL: <http://ochanbhancine.files.wordpress.com/2009/10/butir-soal.doc>. *Butir Soal*. [diakses tanggal 3 Oktober 2011].
- [2] Asmin S., 2008. Pemodelan Nilai UNAS IPA dengan Pendekatan Regresi Semiparametrik Spline di SMAN 1 Grati Pasuruan. *Tesis*. Surabaya, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember.
- [3] Baker F.B., 2001. *The Basics of Item Response Theory. 2nd ed.* ERIC Clearinghouse on Assessment and Evaluation, USA.
- [4] Hambleton R.K., Swaminathan H., & Rogers H.J., 1991. *Fundamental of Item Response Theory*. Sage Publication Inc., Newbury Park, CA.
- [5] Hidayati K., 2002. http://staff.uny.ac.id/sites/default/files/penelitian/Kana_Hidayati_M.Pd./Penerapan_IRT_dalam_Penyetaraan.pdf. *Penerapan IRT dalam Penyetaraan*. [diakses tanggal 22 Agustus 2011].

- [6] Misbah., 2011. <http://id.shvoong.com/social-sciences/education/2190347-pengertian-analisis-item/>. *Pengertian Analisis Item*. [diakses tanggal 25 September 2011].
- [7] Rahmat., 2011. <http://gurupembaharu.com/home/?p=4917>. *Panduan Analisis Butir Soal*. [diakses tanggal 25 September 2011].
- [8] Sehar R., 2000. <http://pustaka.ut.ac.id/pdfpenelitian/70068.pdf>. *Penyetaraan Perangkat Tes Matematika Program D2 PGSD Universitas Terbuka*. [diakses tanggal 5 Agustus 2011].
- [9] Suryabrata S., 1987. *Pengembangan Tes Hasil Belajar*. Jakarta: CV. Raja Wali.
- [10] Vitaria N., 2010. http://eprints.undip.ac.id/22974/1/COVER-ABSTRAK-BAB_I.pdf. *Teori Respon Item Dengan Pendekatan Model Logistik Satu Parameter*. [diakses tanggal 5 Agustus 2011].