



**TEORI RESPON *ITEM* DENGAN MENGGUNAKAN
*RASCH MODEL (RM)***

SKRIPSI

Oleh

**Alfa Rijal Alfian
NIM 071810101108**

**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS JEMBER
2015**



**TEORI RESPON *ITEM* DENGAN MENGGUNAKAN
*RASCH MODEL (RM)***

SKRIPSI

diajukan guna melengkapi tugas akhir dan memenuhi salah satu syarat
untuk menyelesaikan Program Studi Matematika (S1)
dan mencapai gelar Sarjana Sains

Oleh

**Alfa Rijal Alfian
NIM 071810101108**

**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS JEMBER
2015**

PERSEMBAHAN

Skripsi ini saya persembahkan untuk :

1. kedua orang tua tercinta, Ayahanda Moh Masykur dan Ibunda Ida Ayu Titin Aini, terimakasih banyak atas doa, kasih sayang tanpa batas, perhatian, dan segala kebaikan yang telah diberikan, semoga Allah selalu mendekap erat dengan kasih sayang-Nya;
2. kakak-kakakku Dana Ginan Saputra dan Ibnu Fikri Prasetya yang selalu memberikan motivasi tanpa henti;
3. guru-guru sejak taman kanak-kanak hingga perguruan tinggi, yang telah memberikan ilmu serta bimbingan dengan penuh kesabaran;
4. almamater Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember.

MOTTO

Rasulullah SAW. Bersabda, “janganlah kalian menuntut ilmu untuk memanggakannya terhadap para ulama” dan untuk diperdebatkan di kalangan orang-orang bodoh dan buruk perangainya. Jangan pula menuntut ilmu untuk penampilan dalam majelis (pertemuan atau rapat) dan untuk menarik perhatian orang-orang kepadamu. Barangsiapa seperti itu maka baginya neraka...neraka”.

(HR. Attirmidzi dan Ibnu Majah)¹⁾

Manusia, seribu orang di antara mereka seperti satu orang
Seseorang seperti seribu orang apabila memiliki keahlian.

(„Aidh Al-Qarni)²⁾

¹⁾Gus AA. 2009. *Matematika AlQur'an (Mengungkap M'jizat dengan Bahasa Angka)*. Surakarta: Rahma Media Pustaka

²⁾Alqarni, „Aidh. 2010. *Detik-detik Usia yang Paling Mahal*. Yogyakarta: Bening

PERNYATAAN

Saya yang bertanda tangan di bawah ini:

nama : Alfa Rijal Alfian

NIM : 071810101108

menyatakan dengan sesungguhnya bahwa karya ilmiah yang berjudul “Teori Respon *Item* dengan Menggunakan *Rasch Model (RM)*” adalah benar-benar hasil karya sendiri, kecuali kutipan yang sudah saya sebutkan sumbernya, belum pernah diajukan pada institusi mana pun, dan bukan karya jiplakan. Saya bertanggung jawab atas keabsahan dan kebenaran isinya sesuai dengan sikap ilmiah yang harus dijunjung tinggi.

Demikian pernyataan ini saya buat dengan sebenarnya, tanpa ada tekanan dan paksaan dari pihak mana pun serta bersedia mendapat sanksi akademik jika ternyata di kemudian hari pernyataan ini tidak benar.

Jember, Januari 2015

Yang menyatakan,

Alfa Rijal Alfian

NIM 071810101108

SKRIPSI

**TEORI RESPON *ITEM* DENGAN MENGGUNAKAN
*RASCH MODEL (RM)***

Oleh

Alfa Rijal Alfian
NIM 071810101108

Pembimbing

Dosen Pembimbing Utama : Dr. Alfian Futuhul Hadi, S.Si., M.Si

Dosen Pembimbing Anggota : Dian Anggraeni, S.Si., M.Si

PENGESAHAN

Skripsi yang berjudul “Teori Respon *Item* dengan Menggunakan *Rasch Model (RM)*”

telah diuji dan disahkan pada:

hari, tanggal :

tempat : Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember

Tim Penguji:

Ketua,

Sekretaris,

Dr. Alfian Futuhul Hadi, S.Si., M.Si
NIP. 19740719 200012 1 001

Dian Anggraeni, S.Si., M.Si
NIP. 19820216 200604 2 002

Penguji I,

Penguji II,

Prof. Drs. I Made Tirta, M.Sc., Ph.D
NIP. 19591220 198503 1 002

M. Ziaul Arif, S.Si., M.Sc
NIP. 19850111 200812 1 002

Mengesahkan

Dekan,

Prof. Drs. Kusno, DEA., Ph.D.
NIP 19610108 198602 1 001

RINGKASAN

Teori Respon Item Dengan Menggunakan Rasch Model (RM). Alfa Rijal Alfian, 071810101108; 2015: 34 halaman; Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember.

Teori respon *item* merupakan teori pengukuran modern yang digunakan dalam menganalisis *item*. Pada teori respon *item* ada beberapa model yang bisa digunakan dalam analisis, yaitu model logistik satu parameter (*1-PL-model*) adalah model yang melibatkan parameter tingkat kesulitan *item*, model logistik satu parameter biasa disebut *Rasch Model (RM)*. Model logistik dua parameter (*2-PL-model*) adalah model yang melibatkan parameter tingkat kesulitan *item* dan parameter daya beda. Model logistik tiga parameter (*3-PL-model*) adalah model yang melibatkan parameter tingkat kesulitan *item*, daya beda dan parameter tebakan semu (*pseudoguessing*).

Pada penelitian ini model yang digunakan adalah model logistik satu parameter (*1-PL-model*) atau yang biasa disebut *Rasch Model (RM)*. Tujuan penelitian untuk mengetahui penggunaan teori pengukuran modern yaitu teori respon *item* dengan menggunakan *rasch model (RM)* dalam menentukan tingkat kesukaran dan karakteristik *item*. Data yang akan digunakan dalam penelitian ini adalah respon siswa kelas IX Rumah Belajar Genius di Bangil-Pasuruan terhadap materi soal Ujian Nasional mata pelajaran Matematika tahun ajaran 2013/2014. *Item* tes pada penelitian ini dipilih 5 *item* soal secara acak pada materi kelas IX, yaitu *item* 1 merupakan materi tentang kesebangunan, *item* 2 materi tentang akar dan pangkat, *item* 3 materi tentang barisan dan deret, *item* 4 materi tentang peluang dan *item* 5 materi tentang bangun ruang. Analisis menggunakan *software* R dengan paket *ltm*.

Berdasarkan analisis yang telah dilakukan diperoleh tingkat kesukaran *item* paling besar adalah *item* 3 yaitu materi tes tentang barisan dan deret dengan tingkat kesukaran *item* sebesar 1,93; *item* 1 yaitu materi tes tentang kesebangunan dengan tingkat kesukaran *item* sebesar 1,65; *item* 5 yaitu materi tes tentang bangun ruang

dengan tingkat kesukaran *item* sebesar -0,24; *item* 2 yaitu materi tes tentang akar dan pangkat dengan tingkat kesukaran sebesar -0,88 dan tingkat kesukaran paling rendah adalah *item* 4 yaitu materi tes tentang peluang dengan tingkat kesukaran *item* sebesar -1,74.

Berdasarkan fungsi informasi tes yang diperoleh perangkat tes mampu membedakan peserta yang memiliki kemampuan jauh dibawah rata-rata maupun peserta yang memiliki kemampuan jauh diatas rata-rata dengan total informasi yang diberikan oleh perangkat tes tersebut sebesar 4,99. Informasi yang diberikan ketika kemampuan peserta berada jauh dibawah rata-rata ($\theta = -4$) sampai kemampuan peserta berada pada tingkat rata-rata ($\theta = 0$) sebesar 2,24 atau 44,79% dan informasi yang diberikan ketika kemampuan peserta berada pada tingkat rata-rata ($\theta = 0$) sampai kemampuan peserta jauh diatas rata-rata ($\theta = 4$) sebesar 2,36 atau 47,39%.

PRAKATA

Puji syukur ke hadirat Allah SWT atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Teori Respon *Item* dengan Menggunakan *Rasch Model (RM)*”. Skripsi ini disusun untuk memenuhi salah satu syarat menyelesaikan pendidikan strata satu (S1) pada Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember.

Penulisan skripsi ini tidak lepas dari bantuan berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada:

1. Dr. Alfian Futuhul Hadi, S.Si., M.Si selaku Dosen Pembimbing Utama, Dian Anggraeni, S.Si., M.Si selaku Dosen Pembimbing Anggota yang telah meluangkan waktu, pikiran, dan perhatian dalam penulisan skripsi ini;
2. Prof. Drs. I Made Tirta, M.Sc., Ph.D dan M. Ziaul Arif, S.Si., M.Sc selaku dosen penguji yang telah memberi masukan dalam skripsi ini;
3. sahabat Miswandi, Zainal Abidin, Yulan Isa Puspita, Wika Anggani serta semua teman-teman angkatan 2007 jurusan matematika yang telah memberikan dukungan dan semangat untuk menyelesaikan tugas akhir ini;
4. sahabat Haeruddin, Miftahul Ulum, Miftahus Sholihin, Dimas Bagus CW serta sahabat-sahabat Rayon PMII FMIPA lainnya yang tidak bisa disebutkan satu persatu terima kasih atas motivasinya;
5. Sahabat Agus Kiswanto, Mujib Ridwan, Wafiq Umam, Halimah Sa'diyah, Tri Susilo Pramono serta seluruh keluarga besar Rumah Belajar Genius terima kasih atas kerjasamanya dan semua pihak yang tidak dapat disebutkan satu per satu.

Penulis juga menerima segala kritik dan saran dari semua pihak demi kesempurnaan skripsi ini. Akhirnya penulis berharap, semoga skripsi ini dapat bermanfaat.

Jember, Januari 2015

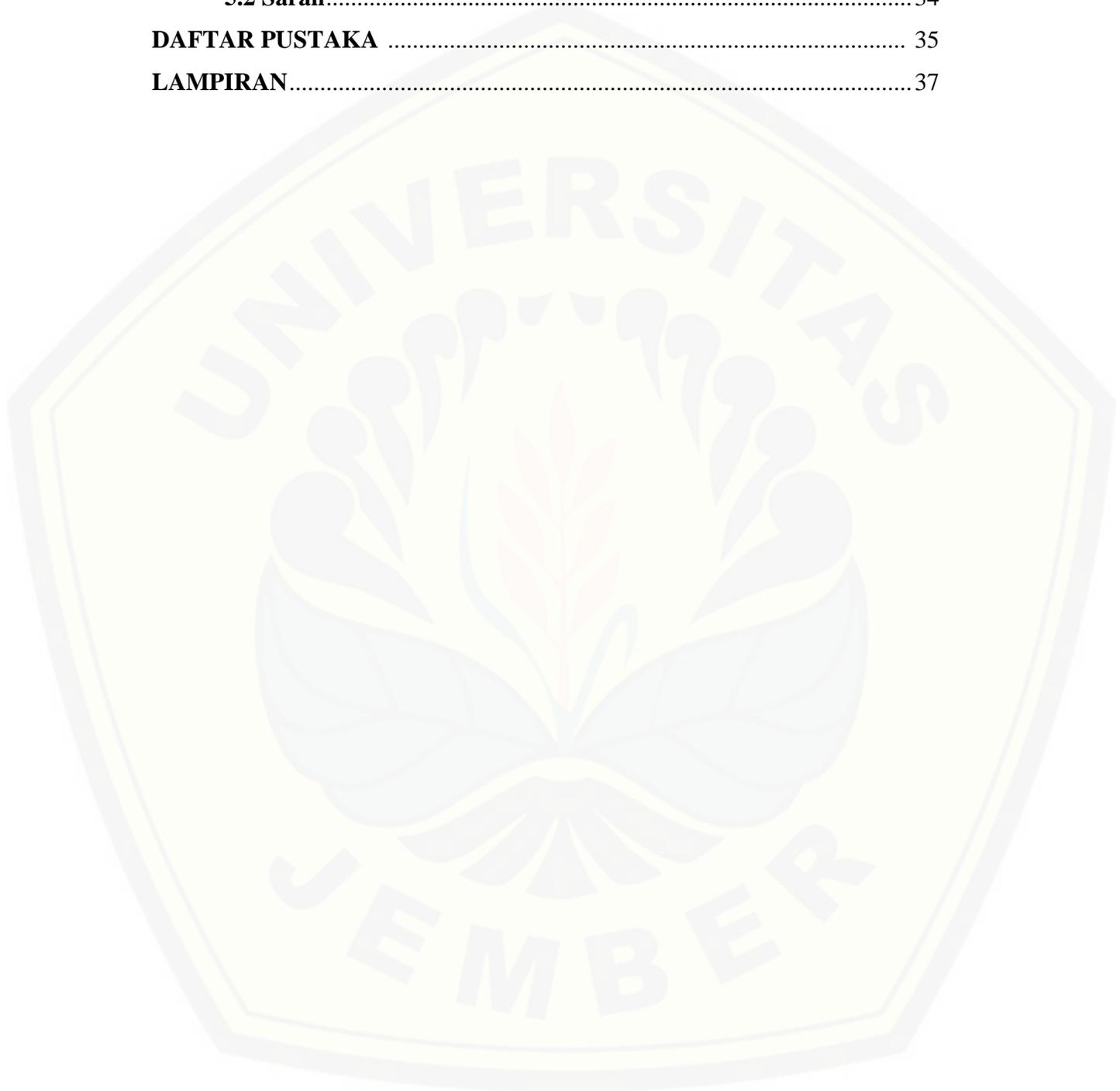
Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PERSEMBAHAN	ii
HALAMAN MOTTO	iii
HALAMAN PERNYATAAN	iv
HALAMAN BIMBINGAN	v
HALAMAN PENGESAHAN	vi
RINGKASAN	vii
PRAKATA	ix
DAFTAR ISI	x
DAFTAR GAMBAR	v
DAFTAR TABEL	vi
DAFTAR LAMPIRAN	vii
BAB 1. PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	4
1.3 Tujuan	4
1.4 Manfaat	4
BAB 2. TINJAUAN PUSTAKA	5
2.1 Teori Uji Klasik	5
2.2 Teori Respon Item	8
2.2.1 Kurva Karakteristik <i>Item</i>	8
2.2.2 Kriteria <i>Item</i> Tes.....	9

2.3 Persyaratan dan Hakikat pada Teori Respon <i>Item</i>	11
2.3.1 Unidimensi	11
2.3.2 Parameter	11
2.3.3 Subpopulasi	12
2.3.4 Independensi Lokal	12
2.3.5 Invariansi Parameter	12
2.4 Model Teori Respon <i>Item</i>	13
2.4.1 Jenis-jenis Teori Respon <i>Item</i>	13
2.4.2 Fungsi Informasi <i>Item</i> dan Fungsi Informasi Tes	14
2.5 Pendugaan Parameter	15
2.6 Kesesuain Model	16
2.7 Pemilihan Model	17
BAB 3. METODE PENELITIAN	18
3.1 Data	18
3.2 Metode Analisis Data	18
3.2.1 Paket Ltm (<i>Latent models</i>)	19
3.2.2 Struktur Fungsi Ltm pada R	19
3.2.3 Langkah-langkah Penelitian	19
BAB 4. HASIL DAN PEMBAHASAN	21
4.1 Gambaran Data	21
4.2 Deskripsi Data	21
4.3 Estimasi Parameter dalam Rasch Model	22
4.4 Kesesuain Model	23
4.5 Penentuan Model	26
4.6 Interpretasi Model	27
4.6.1 Tingkat Kesukaran	27
4.6.2 Kurva Karakteristik <i>Item</i>	28
4.6.3 Fungsi Informasi <i>Item</i> dan Fungsi Informasi Tes	30
BAB 5. PENUTUP	33

5.1 Kesimpulan	33
5.2 Saran.....	34
DAFTAR PUSTAKA	35
LAMPIRAN.....	37



DAFTAR GAMBAR

	Halaman
2.1. Kurva karakteristik <i>item</i>	9
3.1. Langkah-langkah Penelitian	19
4.1. Kurva karakteristik <i>item</i> tes	29
4.2. Kurva fungsi informasi <i>item</i>	30
4.3. Kurva fungsi informasi tes	31

DAFTAR TABEL

	Halaman
4.1. Perbandingan Respon.....	21
4.2. Frekuensi Respon.....	22
4.3. Nilai Parameter Tingkat Kesukaran.....	23
4.4. Uji Dua Arah Model <i>Constrained</i> dan <i>Unconstrained</i>	24
4.5. Uji Tiga Arah Model <i>Constrained</i>	24
4.6. Uji Tiga Arah Model <i>Unconstrained</i>	25
4.7. Nilai AIC dan BIC Model <i>Constrained</i> dan <i>Unconstrained</i>	26
4.8. Nilai AIC dan BIC Model <i>Constrained</i> , 2-PL dan 3-PL.....	27
4.9. Peluang Peserta Menjawab Benar.....	27

DAFTAR LAMPIRAN

	Halaman
A. <i>Item</i> Tes	37
B. Respon Peserta	39
C. Skrip Program untuk <i>Rasch</i> Model Menggunakan <i>Software R</i>	42

BAB 1. PENDAHULUAN

1.1 Latar Belakang

Banyak metode dalam bidang psikometrika dan pendidikan yang digunakan untuk menganalisis *item* dari pendekatan klasik hingga modern. Pendekatan secara klasik yang digunakan adalah teori tes klasik (*classical test theory*) sedangkan pendekatan modern dengan teori respon *item* (*item response theory*).

Teori tes klasik telah lama menjadi acuan dalam pembuatan alat ukur atau instrumen di dunia pendidikan. Dalam teori ini setiap *item* ditelaah menurut tingkat kesukaran *item* dan daya beda *item*. Namun, karakteristik *item* tersebut tergantung pada kelompok sampel mana yang akan digunakan atau dikenai tes. Begitu juga pada koefisien reliabilitas dan validitas tes akan menjadi lebih tinggi apabila kelompok subyek merupakan kelompok yang kemampuannya heterogen (bervariasi besar) dan tes yang sama akan mempunyai koefisien yang lebih rendah apabila dikenakan pada kelompok yang kemampuannya relative homogen. Ketergantungan pada kelompok subyek ini tentu akan mengurangi manfaat parameter *item* dalam berbagai aplikasinya. Kelemahan lain, dalam teori tes klasik diperlukan asumsi kesetaraan *error* pengukuran bagi semua subyek yang dikenai tes. Keberatan asumsi ini adalah kurangnya dukungan yang memperkuatnya dikarenakan pada tes yang sulit, *error* pengukuran bagi subyek yang berkemampuan rendah akan berbeda dari *error* bagi subyek yang berkemampuan tinggi.

Teori respon *item* merupakan teori pengukuran modern yang digunakan dalam menganalisis *item*. Teori ini mempunyai orientasi pada *item* yang karakteristiknya tidak tergantung pada kelompok tertentu. Teori respon *item* 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, dan lain-lain (Nugraheni, 2010).

Pada teori respon *item* ada beberapa model yang bisa digunakan dalam analisis data. Model penskoran dikotomus dimana model respon *items* dikotomus mempunyai dua kategori skor jawaban, yaitu jawaban benar atau tuntas (skor 1) dan jawaban salah atau tidak tuntas (skor 0). Model-model logistik dikotomus dinamai sesuai dengan jumlah parameter yang dilibatkan dalam model. Model logistik satu parameter (*1-PL-model*) adalah model yang melibatkan parameter tingkat kesulitan *item*. Model logistik satu parameter biasa disebut *Rasch Model (RM)*. Model logistik dua parameter (*2-PL-model*) adalah model yang melibatkan parameter tingkat kesulitan *item* dan parameter daya beda. Model logistik tiga parameter (*3-PL-model*) adalah model yang melibatkan parameter tingkat kesulitan *item*, daya beda dan parameter tebakan semu (*pseudoguessing*). Model penskoran politomus adalah model respons *items* yang mempunyai kemungkinan jawaban lebih dari dua kategori. Model penskoran meliputi *Graded Response Model (GRM)* dari Samejima, *Nominal Model (NM)* dari Bock, *Partial Credit Model (PCM)* dari Masters, *Generalized Partial Credit Model (GPCM)* dari Muraki, dan *Rating Scale Model (RSM)* dari Andrich.

Teori respon *item* telah banyak digunakan dalam berbagai penelitian yang berkenaan dengan dunia pendidikan. Pada tahun 2008, Kartono melakukan penelitian untuk mengetahui signifikansi dan kualifikasi level pada faktor-faktor banyaknya parameter *item*, banyaknya kategori *item* politomus, ukuran sampel, dan metode penyetaraan dalam mempengaruhi hasil penyetaraan tes model campuran *3-PL/GPCM* pada tes prestasi belajar yang biasa dilaksanakan pada akhir semester disebut Ulangan Akhir Semester (UAS) untuk mata pelajaran Matematika SMA Negeri dan Swasta kelas XI di kota Semarang. Bentuk tes yang digunakan dalam penelitian tersebut berupa campuran terdiri dari 30 *item* pilihan ganda dengan 5 pilihan jawaban dan 5 *item* uraian dengan 5 kategori tiap *item*. Dalam penelitian tersebut, Kartono (2008) menyimpulkan bahwas faktor-faktor banyaknya parameter

item, banyaknya kategori *item*, ukuran sampel, dan metode penyetaraan yang digunakan berpengaruh pada hasil penyetaraan tes model *3-PL/GPCM*.

Disamping itu, Isgiyanto (2011) melakukan analisis data Ujian Nasional Matematika SMP di wilayah Kabupaten Bantul, Daerah Istimewa Yogyakarta Tahun Pelajaran 2007/2008. Jika Kartono (2008) menggunakan model *3-PL/GPCM*, Isgiyanto (2011) menggunakan *Rasch Model (RM)* dan *Partial Credit Model (PCM)*. Dalam penelitian tersebut, Isgiyanto (2011) ingin menemukan atribut-atribut yang mendasari *item* soal Ujian Nasional Matematika, ketuntasan atribut yang mendasari *item* soal Ujian Nasional Matematika, kemampuan peserta dalam matematika.

Berdasarkan uraian diatas, dapat diketahui bahwa teori respon *item* merupakan analisis yang dibutuhkan pada pengukuran modern hingga saat ini. Dalam penelitian ini, teori respon *item* yang akan digunakan adalah *Rasch Model*. *Rasch model* merupakan analisis teori respon *item* yang menitikberatkan pada parameter tingkat kesukaran. Data yang akan digunakan dalam penelitian ini adalah respon siswa kelas IX Rumah Belajar Genius di Bangil-Pasuruan terhadap materi soal Ujian Nasional mata pelajaran matematika tahun ajaran 2013/2014.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, permasalahan yang akan dibahas pada penelitian ini adalah:

- a. Bagaiman analisis *item* dengan menggunakan teori pengukuran modern yaitu teori respon *item* dengan pendekatan model satu parameter logistik atau *rasch model (RM)*?
- b. Bagaimana penggunaan *rasch model* pada suatu data respons *item* tes dengan model penskoran dikotomus?

1.3 Batasan Masalah

Masalah yang akan dibahas hanya dibatasi pada beberapa hal, yaitu:

- a. Analisis data dilakukan dengan menggunakan teori pengukuran modern yaitu teori respon *item* dengan menggunakan *Rasch Model (RM)*.
- b. Materi soal yang akan digunakan hanya 5 *item* dari 40 *item* yang tersedia pada Ujian Nasional mata pelajaran Matematika dan dipilih secara acak pada materi soal kelas IX.

1.4 Tujuan

Adapun tujuan dari penelitian ini adalah untuk mengetahui penggunaan teori pengukuran modern yaitu teori respon *item* dengan menggunakan *rasch model (RM)* dalam menentukan tingkat kesukaran dan karakteristik *item* pada suatu data respon siswa kelas IX Rumah Belajar Genius di Bangil-Pasuruan terhadap materi soal Ujian Nasional mata pelajaran matematika tahun ajaran 2013/2014.

1.5 Manfaat

Adapun manfaat yang bisa didapatkan dari penelitian ini adalah bermanfaat sebagai sumbangan informasi dalam menganalisis *item* untuk melihat karakteristik *item* sebuah tes dan dapat menambah wawasan serta pengetahuan tentang analisis statistika yang berkaitan dengan teori pengukuran modern yaitu teori respon *item* dengan menggunakan *rasch model (RM)*.

BAB 2. TINJAUAN PUSTAKA

2.1 Teori Uji Klasik

Teori uji klasik merupakan salah satu masalah yang disampaikan oleh pakar psikologi asal Belanda yaitu Charles Spearman dengan menggunakan konsep korelasi (Crocker & Algina, 1986). Ada beberapa aspek yang perlu diperhatikan dalam masalah teori uji klasik yaitu, tingkat kesukaran suatu *item*, daya pembeda *item*, penyebaran pilihan jawaban, dan realibilitas skor tes (Safari, 2009).

a. Tingkat Kesukaran *Item* (p)

Tingkat kesukaran *item* yang dimaksud adalah peluang untuk menjawab dengan benar suatu *item* tes pada tingkat kesukaran tertentu yang biasanya dinyatakan dalam bentuk indeks. Indeks tingkat kesukaran tersebut umumnya dinyatakan dalam proporsi yang besarnya antara 0,00 – 1,00. Rumus yang digunakan untuk menghitung tingkat kesukaran *item* tes pilihan ganda dapat dinyatakan sebagai berikut (Nitko, 1996).

$$p = \frac{\text{Jumlah peserta yang menjawab benar}}{\text{Jumlah seluruh peserta tes}}$$

Berdasarkan nilai p yaitu proporsi peserta tes dalam menjawab *item* tes yang dihasilkan dapat diketahui tingkat kesukaran *item* tes tersebut. Semakin besar nilai p yang berarti semakin besar proporsi peserta tes dalam menjawab benar, maka *item* tes tersebut dianggap mudah. Semakin kecil nilai p maka *item* tes tersebut dianggap sukar.

Tingkat kesukaran *item* tes pada teori uji klasik bergantung kepada kemampuan peserta tes. Pada *item* tes yang mudah akan tampak kemampuan peserta tes menjadi tinggi, namun pada *item* tes yang sukar akan tampak kemampuan peserta tes menjadi rendah. Sedangkan untuk peserta tes yang mempunyai kemampuan yang tinggi, suatu *item* tes akan terasa mudah dan untuk

peserta tes yang mempunyai kemampuan yang rendah, maka *item* tes akan terasa sukar.

Tingkat kesukaran *item* tes diklasifikasikan dalam Nitko (1996) sebagai berikut.

- a. *Item* tes tergolong sukar jika nilai p di antara 0,00 – 0,30,
- b. *Item* tes tergolong sedang jika nilai p di antara 0,31 – 0,70,
- c. *Item* tes tergolong mudah jika nilai p di antara 0,70 – 1,00.

b. Daya Pembeda (DP)

Daya pembeda *item* tes merupakan indeks yang menunjukkan kemampuan *item* tes untuk membedakan kelompok peserta tes yang berkemampuan tinggi dengan peserta tes yang berkemampuan rendah. Daya pembeda suatu *item* tes didasarkan pada hasil suatu tes suatu kelompok sehingga daya pembeda tersebut belum tentu berlaku pada kelompok yang lain. Indeks daya pembeda berkisar antara -1,00 sampai dengan 1,00. Semakin tinggi nilai daya pembeda *item* tes, maka semakin baik *item* tes tersebut.

Daya pembeda *item* tes untuk pilihan ganda dapat dihitung dengan menggunakan rumus berikut.

$$DP = \frac{2(BA - BB)}{N} \quad (2.1)$$

Keterangan:

DP = daya pembeda *item*,

BA = jumlah jawaban benar pada kelompok atas,

BB = jumlah jawaban salah pada kelompok bawah,

N = jumlah peserta tes.

pembagian kelompok atas dan kelompok bawah dalam Anastasi & Urbina (1997) secara umum berdasarkan persentase 25% - 33% urutan nilai terbaik sebagai kelompok atas dan urutan terendah sebagai kelompok bawah.

Daya pembeda *item* tes juga dapat diketahui dengan menggunakan rumus korelasi point biserial (r_{pb}) (Crocker & Algina, 1986). Semakin tinggi nilai

korelasi point biserial, maka semakin baik *item* tes tersebut. Rumus korelasi point biserial dinyatakan sebagai berikut.

$$r_{pb} = \left(\frac{\mu_+ - \mu}{\sigma} \right) \sqrt{\frac{p}{q}} \quad (2.2)$$

Keterangan:

r_{pb} = koefisien point biserial,

μ_+ = rata-rata skor peserta tes yang menjawab benar,

μ = rata-rata skor peserta tes,

σ = simpangan baku skor total,

p = proporsi jawaban benar terhadap semua jawaban peserta,

q = proporsi jawaban salah terhadap semua jawaban peserta.

c. Penyebaran Pilihan Jawaban

Penyebaran pilihan jawaban dijadikan dasar dalam penelaahan *item* tes. Hal ini dimaksudkan untuk mengetahui berfungsi atau tidaknya jawaban yang telah tersedia. Safari (2009) menyatakan bahwa suatu pilihan jawaban (pengecoh) dapat dikatakan berfungsi apabila:

1. Dipilih minimal oleh 5% peserta tes, dan
2. Lebih banyak dipilih oleh kelompok peserta yang belum paham materi.

d. Reliabilitas Skor Tes

Tujuan utama menghitung realibilitas skor tes adalah untuk mengetahui tingkat ketepatan (*precision*) dan kekonsistenan skor tes. Indeks realibilitas berkisar antara 0 – 1. Semakin tinggi koefisien realibilitas suatu tes (mendekati 1), maka semakin tinggi pula ketepatannya.

Realibilitas dapat dihitung dengan koefisien alfa, dalam Crocker & Algina (1986) didefinisikan sebagai berikut:

$$\alpha = \frac{n}{n-1} \left(1 - \frac{\sum_{i=1}^n \sigma_i^2}{\sigma_x^2} \right) \quad (2.3)$$

2.2 Teori Respon *Item*

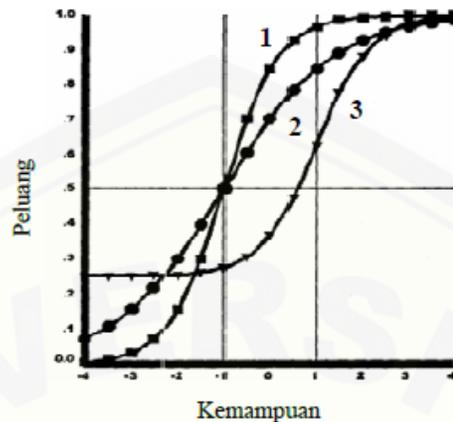
Teori respon *item* dikenal juga sebagai teori ciri laten (*latent trait theory*) atau kelengkungan karakteristik *item* (*item characteristic function*). Teori ini dikembangkan untuk memperbaiki yang terdapat pada uji tes klasik yakni ketergantungan ukuran ciri *item* kepada kelompok peserta serta ketergantungan ukuran ciri peserta kepada kelompok *item* uji tes.

Teori respon *item* membangun suatu model yang menghubungkan ciri *item* dengan ciri peserta. Dengan sejumlah syarat tertentu, model hubungan itu dibuat untuk berlaku secara bebas bagi kelompok *item* dan kelompok peserta mana saja yang memenuhi syarat itu. Sesuai dengan namanya, teori respon *item* terdiri dari tiga unsur. Unsur pertama adalah *item*, unsur kedua adalah peserta yang merespons *item* itu, dan unsur ketiga adalah isi responsi peserta terhadap *item* itu.

Teori respon *item* menelaah *item* untuk menemukan ciri atau karakteristiknya. Setiap penelaahan berkenaan dengan satu *item*, sehingga dalam uji tes didapatkan banyaknya penelaahan, masing-masing berkenaan dengan satu *item* di dalam perangkat uji tes itu. Sedangkan peserta yang merespons *item* itu melakukan responsinya melalui suatu kemampuan. Dalam hal ini, model pada teori respon *item* menggunakan suatu skala kontinu untuk menampung segala jenis kemampuan peserta atau responden yang merespons *item*.

2.2.1 Kurva Karakteristik Item

Kurva karakteristik *item* dalam Hambleton *et al.* (1991) adalah suatu hubungan matematis yang berkaitan dengan peluang berhasil (misalnya, memberikan respon yang benar) pada *item* tes yang diukur dengan melihat kemampuan tes dan karakteristik *item* tes. Semakin tinggi kemampuan seseorang, maka peluang untuk menjawab sebuah *item* tes dengan benar akan semakin meningkat (Crocker & Algina, 1986).



Gambar 2.1. Kurva karakteristik *item*

Keterangan:

item tes 1 ($b=-1;a=1$)

item tes 2 ($b=-1;a=0,5$)

item tes 3 ($b=1;a=1;c=0,25$)

Gambar 2.1 merupakan ilustrasi kurva karakteristik *item* yang diambil dalam Embretson & Reise (2000) terdiri atas tiga *item* tes. Soal nomor 1 dan nomor 2 memiliki tingkat kesukaran yang sama besar yaitu -1, tetapi memiliki nilai daya pembeda yang berbeda yaitu 1 dan 0,5. Perbedaan antara model IRT 1PL dan IRT 2PL terletak pada nilai daya pembeda. Pada model IRT 1PL diasumsikan nilai a (daya pembeda) setiap *item* tes bernilai sama, sedangkan pada model IRT 2PL daya pembeda untuk setiap *item* tes berbeda. Soal nomor 3 merupakan contoh model IRT 3PL dan terdapat peluang menebak sebesar 0,25. Berdasarkan *item* tes nomor 3, dapat dilihat seseorang yang memiliki kemampuan tinggi memiliki peluang untuk menjawab salah dan sebaliknya seseorang yang memiliki kemampuan rendah masih memiliki peluang untuk menjawab benar.

2.2.2 Kriteria *Item* Tes

Pemilihan kriteria *item* tes dilihat berdasarkan uji kesesuaian model dan besarnya nilai parameter *item* tes yaitu daya pembeda, tingkat kesukaran dan peluang

menebak. Menurut Lord (1980) dalam teori respon *item*, *item* yang ideal adalah *item* yang memiliki daya pembeda berkisar 1 (satu) dan peluang menebak mendekati nol. Dalam Hambleton *et al.* (1991), nilai *b* yang mendekati -2,00 menandakan bahwa *item* tes tergolong mudah, dan nilai *b* yang mendekati +2,00 berarti *item* tes tersebut tergolong sukar untuk kelompok tersebut. Berdasarkan hal tersebut maka dapat disimpulkan sebagai berikut:

1. *Item* tes yang baik adalah model setiap *item* yang sesuai dengan model jika nilai khi kuadrat empiris *item* yang tidak melebihi nilai khi kuadrat teoritis (nilai $p > 0,05$), memiliki nilai daya pembeda berkisar 1, peluang menebak kecil, dan tingkat kesukaran diantara nilai -2,00 sampai dengan +2,00.
2. *Item* tes yang cukup baik adalah *item* tes yang sesuai dengan model jika nilai khi kuadrat empiris *item* yang tidak melebihi nilai khi kuadrat teoritis (nilai $p > 0,05$) dan salah satu kriteria *item* tes yang baik tidak terpenuhi.
3. *Item* tes yang belum dapat digambarkan adalah *item* tes yang tidak sesuai dengan model jika nilai khi kuadrat empiris *item* lebih besar dari nilai khi kuadrat teoritis ($p < 0,05$).

Tabel 2.1 Kriteria *item* tes

Model	Kriteria		
	Baik	Cukup Baik	Belum dapat digambarkan
IRT 1PL	Nilai $p > 0,05$; $-2 \leq b \leq 2$	Nilai $p > 0,05$; $b < -2$ atau $b > 2$	Nilai $p < 0,05$
IRT 2PL	Nilai $p > 0,05$; $a \geq 0,5$; $-2 \leq b \leq 2$	Jika salah satu kriteria <i>item</i> tes baik tidak terpenuhi	Nilai $p < 0,05$
IRT 3PL	Nilai $p > 0,05$; $a \geq 0,5$; $-2 \leq b \leq 2$; $c \leq 0,2$	Jika salah satu kriteria <i>item</i> tes baik tidak terpenuhi	Nilai $p < 0,05$

2.3 Persyaratan dan Hakikat pada Teori Respon *Item*

Model pada teori respon *item* itu baru berlaku apabila sejumlah persyaratan dipenuhi. Adapun persyaratan-persyaratan dan hakikat yang berlaku dalam teori respon *item* yaitu unidimensi, parameter, subpopulasi, independensi lokal, invariansi parameter, dan model respon *item*.

2.3.1 Unidimensi

Pada umumnya, teori respon *item* mensyaratkan bahwa setiap *item* hanya mengukur satu macam ciri di kalangan peserta karena setiap ciri ditentukan oleh satu dimensi ukur, maka persyaratan satu macam ciri ini dapat ditafsirkan juga sebagai persyaratan untuk mengukur hanya satu dimensi ciri dari kalangan peserta. Persyaratan *item* unidimensi ini ditujukan untuk mempertahankan invariansi pada teori respon *item*.

2.3.2 Parameter

Ada suatu *item* tertentu yang dijawab oleh sejumlah peserta, jawaban peserta terhadap *item* itu berhubungan dengan ciri *item* dan ciri peserta bersangkutan. Di dalam hubungan ini, ciri peserta dinyatakan melalui parameter ciri peserta θ , ciri *item* dinyatakan dengan tiga parameter *item* a , b , dan c , serta jawaban peserta terhadap *item* itu dinyatakan dalam bentuk probabilitas jawaban benar $P(\theta)$. Dengan demikian, untuk *item* ke- j , hubungan itu dapat ditulis sebagai,

$$P_j(\theta) = \varphi(\theta, a_j, b_j, c_j) \quad (2.4)$$

Di dalam setiap model-model responsi *item* selalu terdapat parameter θ dan probabilitas jawaban $P_j(\theta)$, namun model itu tidak selalu sekaligus memiliki ketiga parameter *item* a_j , b_j , dan c_j . Ada model yang hanya dilengkapi dengan parameter taraf kesukaran *item* b_j dan dikenal sebagai model satu parameter *1-PL*. Ada model yang dilengkapi dengan parameter taraf kesukaran *item* b_j dan parameter daya pembeda *item* a_j dan dikenal sebagai model dua parameter *2-PL*. Dan ada pula model

yang dilengkapi dengan ketiga parameter itu dan dikenal dengan model tiga parameter atau *3-PL*.

2.3.3 Subpopulasi

Parameter ciri peserta θ di dalam model respon *item* membentuk suatu kontinum. Bentangan kontinum ini tidak dibatasi, semuanya bergantung kepada ukuran ciri peserta yang terdapat pada para peserta bersangkutan. Seluruh peserta yang mengerjakan uji tes itu merupakan populasi peserta. Biasanya, teori respon *item* menuntut agar ukuran populasi itu tidaklah terlalu kecil. Pada umumnya, model *2-PL* menuntut ukuran populasi peserta yang lebih besar dari yang dituntut oleh model *1-PL*. Besarnya ukuran populasi ini berkaitan dengan proses estimasi nilai parameter yang kurang cermat, bahkan dalam keadaan tertentu, ukuran populasi yang kecil sama sekali tidak memungkinkan pelaksanaan proses estimasi parameter.

2.3.4 Independensi Lokal

Syarat independensi lokal menentukan bahwa peserta yang di dalam subpopulasi harus independen terhadap *item* uji tes. Dengan independensi lokal, skor dari sejumlah *item* uji tes yang dijawab oleh subpopulasi yang sama, masing-masing haruslah independen. Sehingga skor dari *item* uji tes tidak boleh ditentukan oleh skor pada *item* uji tes yang lain. Independensi lokal adalah independensi secara statistika. Apabila skor untuk N *item* uji tes dihasilkan oleh peserta di dalam subpopulasi itu adalah X_1, X_2, \dots, X_N . Sehingga probabilitas skor P memenuhi ketentuan independensi, jika

$$P(X_1 \cap X_2 \cap \dots \cap X_n) = \prod_{j=1}^N P(X_j) \quad (2.5)$$

2.3.5 Invariansi Parameter

Salah satu syarat pokok dari teori respon *item* adalah invariansi parameter. Melalui syarat ini didapatkan bahwa fungsi atau karakteristik *item* adalah tetap sekalipun kelompok peserta yang menjawab *item* yang sama itu berubah-ubah. Untuk

kelompok yang sama, ciri mereka adalah tetap sekalipun *item* yang mereka jawab itu berubah-ubah.

2.4 Model Teori Respon *Item*

Pada awalnya model yang paling dominan digunakan dalam pengembangan teori respon *item* adalah model ogive normal. Model ogive normal tersebut saat ini sudah jarang digunakan dan digantikan dengan model-model logistik karena model logistik lebih mudah dan sederhana secara prosedur komputasinya dibandingkan dengan model ogive normal.

2.4.1 Jenis-jenis Model Teori Respon *Item*

Saat ini ada tiga model logistik yang sering digunakan (Hambleton *et al.*, 1991) yaitu:

1. Model logistik satu parameter (model *rasch*) atau *item response theory 1-parameter logistic* (IRT 1PL) yaitu untuk menganalisis data yang hanya menitikberatkan pada parameter tingkat kesukaran. Kurva karakteristik *item* tes untuk model satu parameter diberikan oleh persamaan:

$$P_i(\theta) = \frac{e^{(\theta-b_i)}}{1+e^{(\theta-b_i)}} \quad (2.6)$$

2. Model logistik dua parameter atau *item response theory 2-parameter logistic* (IRT 2PL) yaitu untuk menganalisis data yang hanya menitikberatkan pada parameter tingkat kesukaran dan daya pembeda *item* tes. Kurva karakteristik *item* tes untuk model dua parameter diberikan oleh persamaan:

$$P_i(\theta) = \frac{e^{Da_i(\theta-b_i)}}{1+e^{Da_i(\theta-b_i)}} \quad (2.7)$$

3. Model logistik tiga parameter atau *item response theory 3-parameter logistic* (IRT 3PL) yaitu untuk menganalisis data yang menitikberatkan pada parameter tingkat kesukaran *item* tes, daya pembeda *item* tes, dan peluang menebak

(*guessing*). Kurva karakteristik *item* tes untuk model tiga parameter diberikan oleh persamaan:

$$P_i(\theta) = e_i + (1 - e_i) \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad (2.8)$$

Keterangan:

- $P_i(\theta)$ = peluang bahwa peserta tes dengan kemampuan menjawab *item* tes ke-*i* dengan benar,
- a_i = parameter daya pembeda *item* tes ke-*i*,
- b_i = parameter tingkat kesukaran, yaitu satu titik pada skala *ability* dimana kemungkinan untuk menjawab benar sebesar 0,5.
- c_i = peluang tebakan benar *item* tes ke-*i*,
- e = bilangan natural yang nilainya mendekati 2,718
- θ = parameter kemampuan peserta tes,
- D = faktor penskalaan yang diikutkan untuk menjadikan fungsi logistik serupa mungkin dengan fungsi ogive normal ($D = 1,702$).

2.4.2 Fungsi Informasi *Item* dan Fungsi Informasi Tes

Pada teori respon *item*, dikenal fungsi informasi yang digunakan untuk menggambarkan sumbangan/kekuatan *item* soal dalam mengungkap *latent trait*. Sejauhmana masing-masing model tersebut memberikan informasi dijelaskan oleh fungsi informasi atau *information function* (IF) (Veerkamp & Berger, 1999). IF adalah sebuah fungsi sampai sejauh manakah model yang dipilih (1PL, 2PL, atau 3PL) mampu memberikan informasi tentang estimasi *traits-level* sepanjang skala *latent-traits*. Semakin tinggi puncak IF, makin informatif pula model yang dipilih mampu menjelaskan *traits-level* para peserta tes. Oleh karena itu, simpangan baku pengukuran atau *standard error of measurement* (SEM) merupakan fungsi yang berkebalikan dengan IF. Makin tinggi IF, makin rendah SEM. Secara matematis, fungsi informasi *item* diberikan oleh persamaan:

$$I_i(\theta) = P_i(\theta) (1 - P_i(\theta)) \quad (2.9)$$

Sebagai akumulasi dari keseluruhan fungsi informasi *item*, maka akan diperoleh fungsi informasi tes atau *test information function* (TIF), secara matematis diberikan oleh persamaan:

$$TI(\theta) = \sum_{i=1}^n I_i(\theta) \quad (2.10)$$

sedangkan SEM dapat dihitung untuk tiap-tiap kemampuan θ diberikan oleh persamaan:

$$SE(\theta) = \frac{1}{\sqrt{TI(\theta)}} \quad (2.11)$$

2.5 Pendugaan Parameter

Langkah pertama dan paling penting dalam aplikasi teori respon *item* adalah masalah pendugaan parameter, baik parameter kemampuan peserta tes maupun parameter karakteristik *item* (Hambleton *et al.*, 1991). Ada beberapa metode yang digunakan dalam menduga parameter pada model teori respon *item*. Salah satu metode yang digunakan adalah metode kemungkinan maksimum (*maximum likelihood estimation*, MLE).

Prinsip dasar dari metode MLE dalam Hogg *et al.* (1978) adalah jika terdapat contoh acak $X_1, X_2, X_3, \dots, X_n$ dari sebuah sebaran yang memiliki suatu fungsi kepadatan peluang $f(x; \theta), \theta \in \Omega$. Fungsi kepadatan peluang bersama dari $X_1, X_2, X_3, \dots, X_n$ adalah $f(x_1; \theta), f(x_2; \theta), \dots, f(x_n; \theta)$. Fungsi kepadatan peluang bersama ini dipandang sebagai fungsi dari θ . Dengan definisi yang telah dijelaskan maka fungsi nisbah kemungkinan (L) dapat dinotasikan sebagai berikut:

$$L(\theta; x_1, x_2, \dots, x_n) = f(x_1; \theta), f(x_2; \theta), \dots, f(x_n; \theta) \quad (2.9)$$

Hambleton *et al.* (1991) mendefinisikan fungsi likelihood untuk model teori respon *item* sebagai berikut:

$$L(x_1, x_2, \dots, x_n | \theta) = \prod_{j=1}^N \prod_{i=1}^n P_{ij}^{x_{ij}} Q_{ij}^{1-x_{ij}}$$

$$l(X|\theta) = \ln L(x_1, x_2, \dots, x_n | \theta)$$

$$l(X|\Theta) = \ln \left(\prod_{j=1}^N \prod_{i=1}^n P_{ij}^{x_{ij}} Q_{ij}^{1-x_{ij}} \right)$$

$$l(X|\Theta) = \sum_{j=1}^N \sum_{i=1}^n (x_{ij} \ln P_{ij} + (1 - x_{ij}) \ln Q_{ij}) \quad (2.12)$$

dengan

$$\frac{\partial l}{\partial \Theta} = 0, \Theta = (\theta; a, b, c)$$

dimana a , b dan c adalah parameter *item* tes, dan θ adalah parameter kemampuan.

2.6 Kesesuaian Model

Kesesuaian model IRT digunakan untuk menguji karakteristik *item* tes setelah direspon oleh berbagai kemampuan peserta tes. Uji kesesuaian (*goodness of fit test*) model IRT dapat dilakukan untuk setiap *item* tes yang direspon. Pengujian tersebut dilakukan menggunakan statistika Z_{ij}^2 yang berdistribusi khi kuadrat (*chi-square*) (Hambleton *et al.*, 1991) dengan rumus sebagai berikut:

$$Z_{ij} = \frac{P_{ij} - E(P_{ij})}{\sqrt{E(P_{ij})[1 - E(P_{ij})]/N_{ij}}} \quad (2.13)$$

dengan: $\sum_{j=1}^m Z_{ij}^2 \propto \chi_{m-k}^2$

P_{ij} = proporsi jawaban benar pada *item* ke- i dalam kelompok kemampuan ke- j

$E(P_{ij})$ = nilai harapan dari p_{ij}

N_{ij} = jumlah peserta tes yang merespon *item* ke- i dalam kelompok kemampuan ke- j

K = jumlah parameter

m = jumlah kategori kemampuan yang bergantung pada pendugaan kemampuan peserta tes

Hipotesis nol dalam pengujian ini menyatakan bahwa kurva karakteristik sesuai data respon *item* ke- i . Kesesuaian model untuk suatu *item* dapat ditunjukkan oleh nilai khi kuadrat empiris *item* yang tidak melebihi nilai khi kuadrat teoritis.

2.7 Pemilihan Model

Analisis yang digunakan untuk memilih model adalah *likelihood ratio test* (LRT). LRT merupakan suatu uji yang membandingkan nilai fungsi *likelihood* untuk dua model yang akan diuji. Secara statistik uji IRT didefinisikan dalam Wiberg (2007) sebagai berikut:

$$G^2 = -2 \ln \frac{l(a)}{l(c)} = -2 \ln[l(a) - l(c)] \propto \chi^2_{(m)} \quad (2.14)$$

dengan:

$l(a)$ = nilai fungsi *likelihood* dari model *augmented*

$l(c)$ = nilai fungsi *likelihood* dari model *compact*

m = selisih jumlah parameter antara model *augmented* dan model *compact*

Pemilihan model dapat juga dilakukan dengan melihat nilai AIC (*Akaike's Information Criterion*) dan BIC (*Bayesian Information Criterion*). Secara umum, semakin kecil nilai AIC maka model yang dipakai semakin cocok. Model yang dianggap terbaik adalah model dengan nilai AIC minimum. Namun demikian, dengan pertimbangan aspek lain, perbedaan AIC yang tidak terlalu besar mungkin dapat diabaikan (Tirta, 2009).

Selain menggunakan kriteria statistik AIC (*Akaike's Information Criterion*), untuk memilih model yang terbaik bisa digunakan juga BIC (*Bayesian Information Criterion*). *Bayesian Information Criterion* (BIC) adalah sebuah kriteria statistik untuk memilih model. Nilai BIC mencerminkan peningkatan nilai jumlah kuadrat residual dan jumlah parameter dari model yang digunakan. Variasi yang tidak bisa dijelaskan oleh model mengakibatkan peningkatan BIC. Model dengan nilai BIC yang lebih kecil dipilih sebagai model yang terbaik, karena nilai BIC yang lebih kecil menunjukkan bahwa model yang dihasilkan lebih bisa menjelaskan variasi dari data.

BAB 3. METODE PENELITIAN

Pada bab ini akan diuraikan mengenai metode penelitian sebagai konsep dalam melakukan penelitian ini. Metode penelitian ini meliputi data dan analisis data dengan menggunakan *Rasch Model (RM)*.

3.1 Data

Sumber data yang digunakan dalam penelitian ini adalah data primer yang berupa respon siswa kelas IX Rumah Belajar Genius di Bangil-Pasuruan dengan jumlah peserta sebanyak 80 siswa terhadap materi soal Ujian Nasional mata pelajaran Matematika tahun ajaran 2013/2014. Respon siswa akan diskor berdasarkan model penskoran dikotomis, yaitu 0 jika jawaban salah dan 1 jika jawaban benar. Hasil respon dapat dilihat pada Lampiran B.

Soal Ujian Nasional untuk mata pelajaran Matematika terdiri dari 40 *item* soal dengan komposisi 8 *item* soal atau 20% materi kelas VII, 12 *item* soal atau 30% materi kelas VIII dan 20 *item* soal atau 50% materi kelas IX. Pada penelitian ini dipilih 5 *item* soal secara acak pada materi kelas IX, yaitu *item* 1 merupakan materi tentang kesebangunan, *item* 2 materi tentang akar dan pangkat, *item* 3 materi tentang barisan dan deret, *item* 4 materi tentang peluang dan *item* 5 materi tentang bangun ruang. *Item* tes yang digunakan pada penelitian ini dapat dilihat pada Lampiran A.

3.2 Metode Analisis Data

Metode analisis yang digunakan dalam penelitian ini adalah analisis variabel laten, yaitu *rasch model*. Software yang digunakan adalah *software open source* (paket R) versi 3.1.2. Dalam hal ini paket yang dipakai adalah paket *ltm*.

3.2.1 Paket ltm

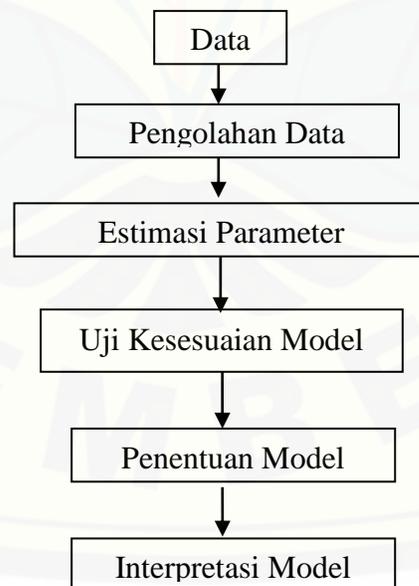
Paket `ltm` dikembangkan untuk menangani masalah analisis multivariat dengan data variabel laten yang bersifat biner dan politomus, dengan menggunakan pendekatan teori respon *item*. Paket `ltm` bisa didapatkan dengan *men-download* di <http://CRAN.R-project.org/>.

3.2.2 Struktur Fungsi ltm pada R

Paket `ltm` menyediakan fungsi untuk dua jenis data. Pertama, apabila jenis data termasuk biner maka struktur fungsi yang dipakai adalah `rasch()`, `ltm()` dan `tpm()`. Kedua, apabila data yang dipakai adalah data politomus maka fungsi yang digunakan adalah `grm()`.

3.2.3 Langkah-langkah penelitian

Langkah-langkah yang akan dilakukan dalam penelitian ini dapat dilihat pada Gambar 3.1



Gambar 3.1 Langkah-langkah Penelitian

Langkah penelitian pada Gambar 3.1 dapat dijelaskan sebagai berikut.

1. Data
Data yang digunakan seperti yang sudah dijelaskan pada subbab 3.1. kemudian data tersebut dimasukan dalam *software* program R 3.1.2
2. Estimasi parameter
Estimasi parameter dilakukan menggunakan metode *maximum likelihood estimation (MLE)*. Parameter yang akan diduga adalah parameter tingkat kesukaran tiap *item*.
3. Uji kesesuaian model
Menentukan model dilakukan dengan menggunakan pengujian statistik Z_{ij}^2 yang berdistribusi khi-kuadrat. Uji dilakukan menggunakan uji dua arah dan uji tiga arah. Dalam hal ini, dicari dua model berdasarkan data yang diketahui. Model yang satu menggunakan *constrained*, sedangkan model yang kedua menggunakan *unconstrained*.
4. Pemilihan model yang cocok
Menentukan model yang cocok bisa dilakukan dengan membandingkan nilai AIC dan BIC pada model kelompok yang satu dengan kelompok yang lain. Semakin kecil nilai AIC atau BIC, maka semakin cocok model yang didapatkan.
5. Interpretasi model yang cocok
Dari model yang didapatkan pada langkah 4, kemudian menginterpretasikan model tersebut dan akan ditunjukkan kurva karakteristik *item* dan fungsi informasi *item* serta fungsi informasi tes.

BAB 4. HASIL DAN PEMBAHASAN

4.1 Gambaran Data

Penelitian ini menggunakan data yang telah dijelaskan pada subbab 3.1. Pada data tersebut akan dilakukan analisis menggunakan teori respon *item* model logistik satu parameter (model *rasch*) atau *item response theory 1-parameter logistic* (IRT 1PL) yang bertujuan untuk mengetahui tingkat kesukaran *item* tes.

4.2 Deskripsi Data

Pada tahapan ini akan dilakukan proses analisis data dengan menggunakan bantuan *software* program R 3.1.2. Ada beberapa hal yang akan dilakukan diantaranya:

a. Perbandingan Respon Peserta Terhadap *Item* Tes

Hasil analisis data terkait perbandingan respon peserta terhadap *item* tes disajikan pada tabel 4.1 dibawah ini

Tabel 4.1 Perbandingan Respon

<i>Item</i> Tes	Perbandingan Respon (%)	
	Respon Salah (0)	Respon Benar (1)
1	80	20
2	33	77
3	84	16
4	19	81
5	45	55

berdasarkan Tabel 4.1 dapat dilihat bahwa respon salah paling besar secara persentase adalah *item* tes 3 dengan persentase sebesar 84% dan respon salah paling kecil secara persentase adalah *item* tes 4 dengan persentase sebesar 19%.

b. Frekuensi Respon Peserta Terhadap *Item* tes

Hasil analisis data terkait frekuensi respon peserta terhadap *item* tes disajikan pada Tabel 4.2 dibawah ini

Tabel 4.2 Frekuensi Respon

Respon Peserta	Frekuensi
0	4
1	14
2	26
3	21
4	12
5	3

Tabel 4.2 menunjukkan bahwa ada 4 peserta yang merespon ke-5 *item* tes dan keseluruhannya salah, 14 peserta yang merespon ke-5 *item* tes dan hanya 1 *item* tes yang benar, 26 peserta yang merespon ke-5 *item* tes dan hanya 2 *item* yang benar, 21 peserta yang merespon ke-5 *item* tes dan hanya 3 *item* yang benar, 12 peserta yang merespon ke-5 *item* tes dan hanya 4 *item* tes yang benar, 3 peserta yang merespon ke-5 *item* tes dan keseluruhannya benar.

4.3 Estimasi Parameter dalam *Rasch* Model

Teori respon *item* model logistik satu parameter (IRT 1PL) atau yang biasa disebut model *rasch* adalah analisis data yang hanya menitikberatkan pada parameter tingkat kesukaran. Estimasi parameter sangat diperlukan untuk mengetahui karakteristik *item* soal maupun kemampuan peserta tes. Terdapat pendugaan parameter pada 2 model *item* tes yaitu, model *constrained* dan *unconstrained*. Hasil 2

model pendugaan parameter karakteristik *item* setiap *item* tes menggunakan model IRT 1-PL (*rasch*) menunjukkan bahwa setiap *item* tes memiliki nilai *b* (tingkat kesukaran) yang cukup beragam. Berikut nilai *b* yang didapatkan dengan bantuan *software* R disajikan pada Tabel 4.3

Tabel 4.3 Nilai Parameter Tingkat Kesukaran

Item Tes	<i>Constrained</i>		<i>Unconstrained</i>	
	<i>b</i>	std. err	<i>B</i>	std. err
1	1,65	0,32	1,61	0,41
2	-0,88	0,28	-0,86	0,31
3	1,94	0,34	1,89	0,46
4	-1,74	0,33	-1,70	0,43
5	-0,24	0,27	-0,24	0,27

Berdasarkan Tabel 4.3 dapat dilihat bahwa parameter tingkat kesukaran untuk model *constrained* berkisar antara -1,74 sampai 1,95 sedangkan untuk model *unconstrained* berkisar antara -1,70 sampai 1,89. *Item* yang dinyatakan baik adalah *item* yang memiliki tingkat kesukaran (*b*) berkisar $-2 \leq b \leq +2$.

4.4 Kesesuaian Model

Kesesuaian model digunakan untuk menguji karakteristik *item* tes setelah direspon oleh berbagai kemampuan peserta tes. Uji kesesuaian (*goodness of fit test*) model dapat dilakukan untuk setiap *item* tes yang direspon dengan uji dua arah maupun uji tiga arah dan melihat nilai dari χ^2 residu. Uji dua arah dan uji tiga arah dilakukan dengan menyusun semua kemungkinan tiap hasil respon yang ada. Uji dua arah untuk model *constrained* dan *unconstrained* dapat dilihat pada Tabel 4.4

Tabel 4.4 Uji Dua Arah Model *Constrained* dan *Unconstrained*

Respon	<i>Constrained</i>					<i>Unconstrained</i>				
	Item <i>i</i>	Item <i>j</i>	Obs	Exp	(O-E) ² /E	Item <i>i</i>	Item <i>j</i>	Obs	Exp	(O-E) ² /E
(0,0)	4	5	11	8,94	0,47	4	5	11	9,08	0,40
	2	5	13	14,75	0,21	2	5	13	14,93	0,25
	2	4	6	6,94	0,13	2	4	6	7,08	0,16
(1,0)	3	5	2	3,92	0,94	3	5	2	3,83	0,87
	1	5	3	4,96	0,77	1	5	3	4,84	0,70
	3	4	2	1,34	0,33	3	4	2	1,29	0,39
(0,1)	4	5	4	6,03	0,68	4	5	4	5,92	0,62
	2	5	13	11,21	0,28	2	5	13	11,05	0,34
	1	3	10	8,88	0,14	1	3	10	8,80	0,16
(1,1)	3	5	11	9,05	0,42	3	5	11	9,17	0,37
	1	5	13	11,02	0,36	1	5	13	4,20	0,34
	1	3	3	4,09	0,29	1	3	3	11,16	0,30

Sedangkan uji tiga arah untuk model *constrained* dan *unconstrained* dapat dilihat pada Tabel 4.5 dan Tabel 4.6

Tabel 4.5 Uji Tiga Arah Model *Constrained*

Respon	Item <i>i</i>	Item <i>j</i>	Item <i>k</i>	Obs	Exp	(O-E) ² /E
(0,0,0)	3	4	5	11	8,35	0,84
	2	3	4	5	6,52	0,35
(1,0,0)	1	4	5	2	0,77	1,99
	2	4	5	7	4,25	1,78
(0,1,0)	1	3	4	2	1,08	0,80
	2	3	4	1	0,42	0,79
(1,1,0)	1	4	5	1	4,19	2,43
	2	3	5	1	2,83	1,19

Tabel 4.5 Uji Tiga Arah Model *Constrained* (lanjutan)

Respon	Item <i>i</i>	Item <i>j</i>	Item <i>k</i>	Obs	Exp	(O-E) ² /E
(0,0,1)	3	4	5	2	5,29	2,05
	1	3	5	23	27,17	0,64
(1,0,1)	3	4	5	2	0,74	2,12
	1	4	5	0	0,94	0,94
(0,1,1)	1	3	5	8	5,84	0,80
	2	4	5	11	8,97	0,46
(1,1,1)	1	4	5	13	10,09	0,84
	2	3	5	9	7,55	0,28

Tabel 4.6 Uji Tiga Arah Model *Unconstrained*

Respon	Item <i>i</i>	Item <i>j</i>	Item <i>k</i>	Obs	Exp	(O-E) ² /E
(0,0,0)	3	4	5	11	8,51	0,73
	2	3	4	5	6,67	0,42
(1,0,0)	1	4	5	2	0,74	2,15
	2	4	5	7	4,22	1,83
(0,1,0)	1	3	4	2	1,04	0,90
	2	3	4	1	0,41	0,87
(1,1,0)	1	4	5	1	4,10	2,35
	2	3	5	1	2,78	1,14
(0,0,1)	3	4	5	2	5,20	1,97
	1	3	5	23	27,04	0,60
(1,0,1)	3	4	5	2	0,72	2,30
	1	4	5	0	0,90	0,90
(0,1,1)	1	3	5	8	5,83	0,81
	2	4	5	11	8,83	0,53
(1,1,1)	1	4	5	13	10,25	0,74
	1	3	4	3	3,94	0,23

Berdasarkan uji kesesuaian model dengan menggunakan uji dua arah maupun uji tiga arah pada model *constrained* dan *unconstrained* terlihat bahwa tidak ada nilai χ^2 residu yang lebih dari 3,5 yang berarti bahwa model *constrained* dan *unconstrained* merupakan model yang sesuai.

4.5 Penentuan Model

Setelah diketahui model yang sesuai kemudian menentukan model yang cocok berdasarkan nilai AIC dan BIC masing-masing model. Semakin kecil nilai tersebut, maka semakin baik model yang didapatkan. Nilai AIC dan BIC untuk model *constrained* dan *unconstrained* masing-masing dapat dilihat pada Tabel 4.7

Tabel 4.7 Nilai AIC dan BIC Model *Constrained* dan *Unconstrained*

Model	AIC	BIC
<i>Constrained</i>	436,77	448,68
<i>Unconstrained</i>	438,74	453,04

Berdasarkan informasi pada Tabel 4.7 tentang nilai AIC dan BIC dapat diketahui bahwa nilai AIC pada model *constrained* sebesar 436,77. Nilai ini lebih kecil bila dibandingkan dengan nilai AIC pada model *unconstrained* yaitu 438,74. Sedangkan untuk nilai BIC pada model *constrained* juga lebih kecil yaitu sebesar 448,68 bila dibandingkan dengan nilai BIC pada model *unconstrained*. Berdasarkan nilai tersebut maka model yang cocok adalah model *constrained*.

Setelah model yang cocok didapatkan yaitu model *constrained*, akan dibandingkan pula dengan model teori respon *item* dua parameter logistik (2-PL) dan model teori respon *item* tiga parameter logistik (3-PL). Perbandingan model tersebut dilakukan dengan cara melihat nilai AIC dan BIC pada masing-masing model pada. Nilai AIC dan BIC yang dihasilkan pada model tersebut disajikan pada Tabel 4.8

Tabel 4.8 Nilai AIC dan BIC Model *Constrained*, 2-PL dan 3-PL

Model	AIC	BIC
<i>Constrained</i>	436,77	448,68
2-PL	441,16	464,98
3-PL	444,61	470,82

Pada Tabel 4.8 dapat dilihat bahwa Nilai AIC pada model *constrained* lebih kecil yaitu sebesar 436,77 bila dibandingkan dengan model 2-PL sebesar 441,16 dan model 3-PL sebesar 444,61. Sedangkan untuk nilai BIC yang didapat pada model *constrained* pun paling kecil yaitu sebesar 448,68. Berdasarkan nilai AIC dan BIC yang didapat pada maka model yang cocok adalah model *constrained*.

4.6 Interpretasi Model

Setelah didapatkan model yang baik (*constrained*) maka di interpretasikan beberapa hal yang ada pada analisis *rasch* model diantaranya adalah tingkat kesukaran dan peluang berhasil (memberikan respons benar) pada *item* tes, kurva karakteristik *item* tes, fungsi informasi *item* dan fungsi informasi tes.

4.6.1 Tingkat Kesukaran

Berdasarkan model yang cocok maka tingkat kesukaran *item* dan peluang berhasil menjawab *item* dengan benar akan disajikan pada Tabel 4.9

Tabel 4.9 Peluang Peserta Menjawab Benar

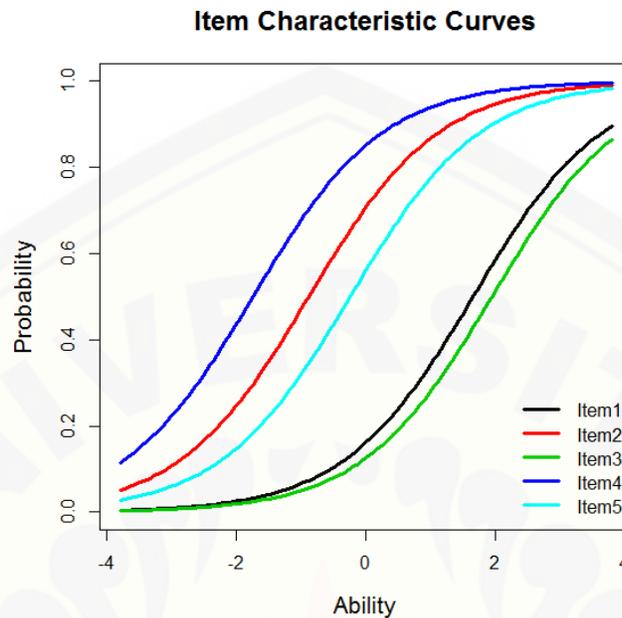
<i>Item</i> Tes	Tingkat Kesukaran (b)	Daya Beda <i>Item</i> (a)	$P_i(\theta)$
4	-1,74	1	0,85
2	-0,88	1	0,71
5	-0,24	1	0,56
1	1,65	1	0,16
3	1,94	1	0,12

Tabel 4.9 menjelaskan bahwa *item* tes dengan tingkat kesukaran paling rendah adalah *item* 4 dengan nilai tingkat kesukaran sebesar -1,74 dan peluang peserta menjawab dengan benar *item* tersebut ketika mempunyai kemampuan rata-rata adalah sebesar 0,85. Sedangkan *item* tes dengan tingkat kesukaran paling tinggi adalah *item* 3 dengan nilai tingkat kesukaran sebesar 1,94 dan peluang peserta menjawab dengan benar *item* tersebut ketika mempunyai kemampuan rata-rata adalah sebesar 0,12.

4.6.2 Kurva Karakteristik *Item*

Kurva karakteristik *item* merupakan hubungan matematis yang berkaitan dengan peluang berhasil (memberikan respons benar) pada *item* tes yang diukur dengan melihat kemampuan tes dan karakteristik *item* tes. Semakin tinggi kemampuan peserta tes maka peluang untuk menjawab sebuah *item* tes dengan benar akan semakin meningkat. Lengkungan kurva model *Rasch* hanya ditentukan oleh parameter tingkat kesukaran *item* b_i dan parameter ciri peserta yang menjawab *item* tersebut. Jadi peluang peserta menjawab benar suatu *item* merupakan fungsi dari kemampuan peserta dan tingkat kesukaran *item*. Tingkat kesukaran *item* tes bergerak dari $-\infty$ sampai $+\infty$. Meski demikian nilai yang berarti biasanya bergerak dari skala -3 sampai +3 dalam satuan logit (log odd unit). *Item* yang memiliki tingkat kesukaran (b_i) dekat atau terletak di bawah skala -2,00 menunjukkan bahwa *item* tes tersebut termasuk kategori mudah. *Item* yang memiliki tingkat kesukaran (b_i) dekat atau terletak di atas skala +2,00 menunjukkan *item* tes tersebut termasuk kategori sukar. *Item* yang dinyatakan baik adalah *item* yang memiliki tingkat kesukaran (b_i) berkisar $-2 \leq b_i \leq +2$.

Berikut kurva karakteristik *item* yang dihasilkan pada penelitian ini untuk kelima *item* tes disajikan dalam Gambar 4.1

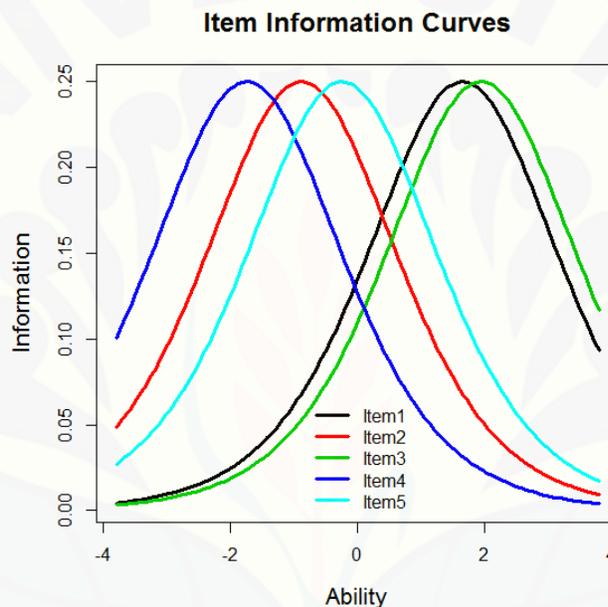


Gambar 4.1 Kurva karakteristik *item* tes

Gambar 4.1 menunjukkan bahwa peluang menjawab dengan benar suatu *item* tes merupakan fungsi dari kemampuan (*latent trait*) peserta tes. Kurva tersebut menunjukkan tingkat kesukaran dari ke-5 *item* tes 1, 2, 3, 4 dan 5. Kelima *item* tes memiliki tingkat kesukaran yang berbeda-beda. Kurva *item* 1 menggambarkan sebuah *item* tes dengan tingkat kesukaran $b= 1,65$; kurva *item* 2 menggambarkan sebuah *item* tes dengan tingkat kesukaran $b= -0,88$; kurva *item* 3 menggambarkan sebuah *item* tes dengan tingkat kesukaran $b= 1,94$; kurva *item* 4 menggambarkan sebuah *item* tes dengan tingkat kesukaran $b= -1,74$; dan kurva *item* 5 menggambarkan sebuah *item* tes dengan tingkat kesukaran $b= -0,24$. Kurva *item* tes pada Gambar 4.1 menggambarkan bahwa peserta tes dengan kemampuan (*latent trait*) rendah akan mempunyai peluang menjawab *item* tes dengan benar sangat kecil, sedangkan peserta tes dengan kemampuan tinggi akan memiliki peluang untuk menjawab dengan benar besar pada *item* tes dengan tingkat kesukaran yang besar. Namun pada *item* tes dengan tingkat kesukaran yang kecil, peserta yang memiliki kemampuan rendah mempunyai peluang menjawab yang besar, begitu juga dengan peserta tes yang memiliki kemampuan yang tinggi.

4.6.3 Fungsi Informasi *Item* dan Fungsi Informasi Tes

Fungsi informasi *item* (*Item Information Functions*) merupakan suatu metode untuk menjelaskan kekuatan suatu *item* pada perangkat tes, pemilihan *item* tes, dan perbandingan beberapa perangkat tes. Fungsi informasi *item* menyatakan informasi *item* tes dalam mengungkap kemampuan peserta (*latent trait*) yang diukur dengan tes tersebut. Semakin tinggi puncak fungsi informasi maka semakin baik pula informasi yang diberikan. Fungsi informasi *item* dapat dilihat pada Gambar 4.2

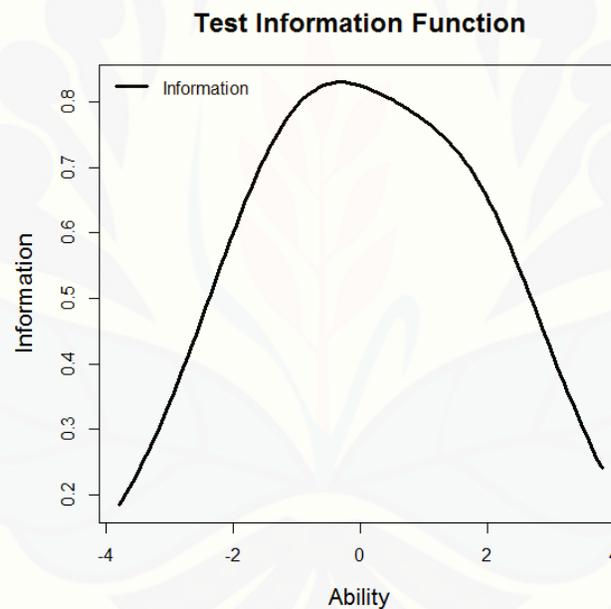


Gambar 4.2 Kurva fungsi informasi *item*

Pada Gambar 4.2 terlihat bahwa fungsi informasi *item* yang yang diberikan beragam dan berbeda satu sama lain. Sebagai contoh untuk *item* 3, memberikan informasi paling tinggi ketika kemampuan peserta berada pada tingkat $\theta = 2$, ketika kemampuan peserta berada pada tingkat diatas rata-rata $\theta = 4$ pun masih mampu memberikan informasi berkisar 0,12; akan tetapi tidak ada informasi yang diberikan ketika kemampuan peserta dibawah rata-rata ($\theta = -4$). Hasil tersebut berarti bahwa *item* 3 lebih mampu membedakan peserta yang memiliki kemampuan tinggi ketimbang peserta yang memiliki kemampuan rendah. Sedangkan untuk *item* 4 berkebalikan dengan *item* 3. *Item* 4 memberikan informasi ketika kemampuan peserta

berada jauh dibawah rata-rata ($\theta = -4$), namun tidak memberikan informasi ketika kemampuan peserta berada jauh diatas rata-rata ($\theta = 4$). *Item 4* lebih mampu membedakan peserta yang memiliki kemampuan rendah ketimbang peserta yang memiliki kemampuan tinggi.

Test Information Function (TIF) merupakan jumlah dari fungsi informasi *item* penyusun tes tersebut. Berhubungan dengan hal ini, fungsi informasi perangkat tes akan tinggi jika butir tes mempunyai fungsi informasi yang tinggi pula. Kurva informasi tes berguna untuk menggambarkan sejauh mana tes menyediakan perbedaan kualitas informasi pada tingkat kemampuan yang berbeda. Pada Gambar 4.3 akan disajikan kurva dari TIF tersebut



Gambar 4.3 Kurva fungsi informasi tes

Pada Gambar 4.3 terlihat bahwa perangkat tes dapat memberikan informasi paling besar pada tingkat peserta yang memiliki kemampuan rata-rata ($\theta = 0$). Perangkat tes mampu membedakan peserta yang memiliki kemampuan jauh dibawah rata-rata ($\theta = -4$) maupun peserta yang memiliki kemampuan jauh diatas rata-rata ($\theta = 4$). Informasi yang diberikan ketika kemampuan peserta berada jauh dibawah rata-rata ($\theta = -4$) sampai kemampuan peserta berada pada tingkat rata-rata ($\theta = 0$) sebesar

2,24 atau 44,79%. Sedangkan Informasi yang diberikan ketika kemampuan peserta berada pada tingkat rata-rata ($\theta = 0$) sampai kemampuan peserta jauh di atas rata-rata ($\theta = 4$) sebesar 2,36 atau 47,39%. Secara keseluruhan total informasi yang diberikan oleh perangkat tes tersebut sebesar 4,99.



BAB 5 PENUTUP

5.1 Kesimpulan

Berdasarkan hasil analisis data pada Bab 4 menggunakan teori respon *item* dengan pendekatan *rasch model*, dapat ditarik kesimpulan sebagai berikut:

1. Berdasarkan nilai AIC dan BIC yang didapat pada model *rasch constrained* secara berturut-turut yaitu sebesar 436,77 dan 448,68. Nilai tersebut paling kecil bila dibandingkan dengan model-model yang ada, maka model *rasch constrained* merupakan model yang cocok.
2. Tingkat kesukaran *item* paling besar adalah *item* 3 yaitu materi tes tentang barisan dan deret dengan tingkat kesukaran *item* sebesar 1,93; *item* 1 yaitu materi tes tentang kesebangunan dengan tingkat kesukaran *item* sebesar 1,65; *item* 5 yaitu materi tes tentang bangun ruang dengan tingkat kesukaran *item* sebesar -0,24; *item* 2 yaitu materi tes tentang akar dan pangkat dengan tingkat kesukaran sebesar -0,88 dan tingkat kesukaran paling rendah adalah *item* 4 yaitu materi tes tentang peluang dengan tingkat kesukaran *item* sebesar -1,74.
3. Berdasarkan fungsi informasi tes yang diperoleh perangkat tes mampu membedakan peserta yang memiliki kemampuan jauh dibawah rata-rata maupun peserta yang memiliki kemampuan jauh diatas rata-rata dengan total informasi yang diberikan oleh perangkat tes tersebut sebesar 4,99. Informasi yang diberikan ketika kemampuan peserta berada jauh dibawah rata-rata ($\theta = -4$) sampai kemampuan peserta berada pada tingkat rata-rata ($\theta = 0$) sebesar 2,24 atau 44,79% dan informasi yang diberikan ketika kemampuan peserta berada pada tingkat rata-rata ($\theta = 0$) sampai kemampuan peserta jauh diatas rata-rata ($\theta = 4$) sebesar 2,36 atau 47,39%.

5.2 Saran

Pada skripsi ini, analisis yang digunakan adalah teori respon *item* dengan menggunakan pendekatan model *rasch* atau model logistik satu parameter (1-PL). Ada kesempatan bagi para peneliti untuk mengembangkan analisis ini, misalnya dengan menggunakan pendekatan model logistik dua parameter (2-PL) ataupun model logistik tiga parameter (3-PL).



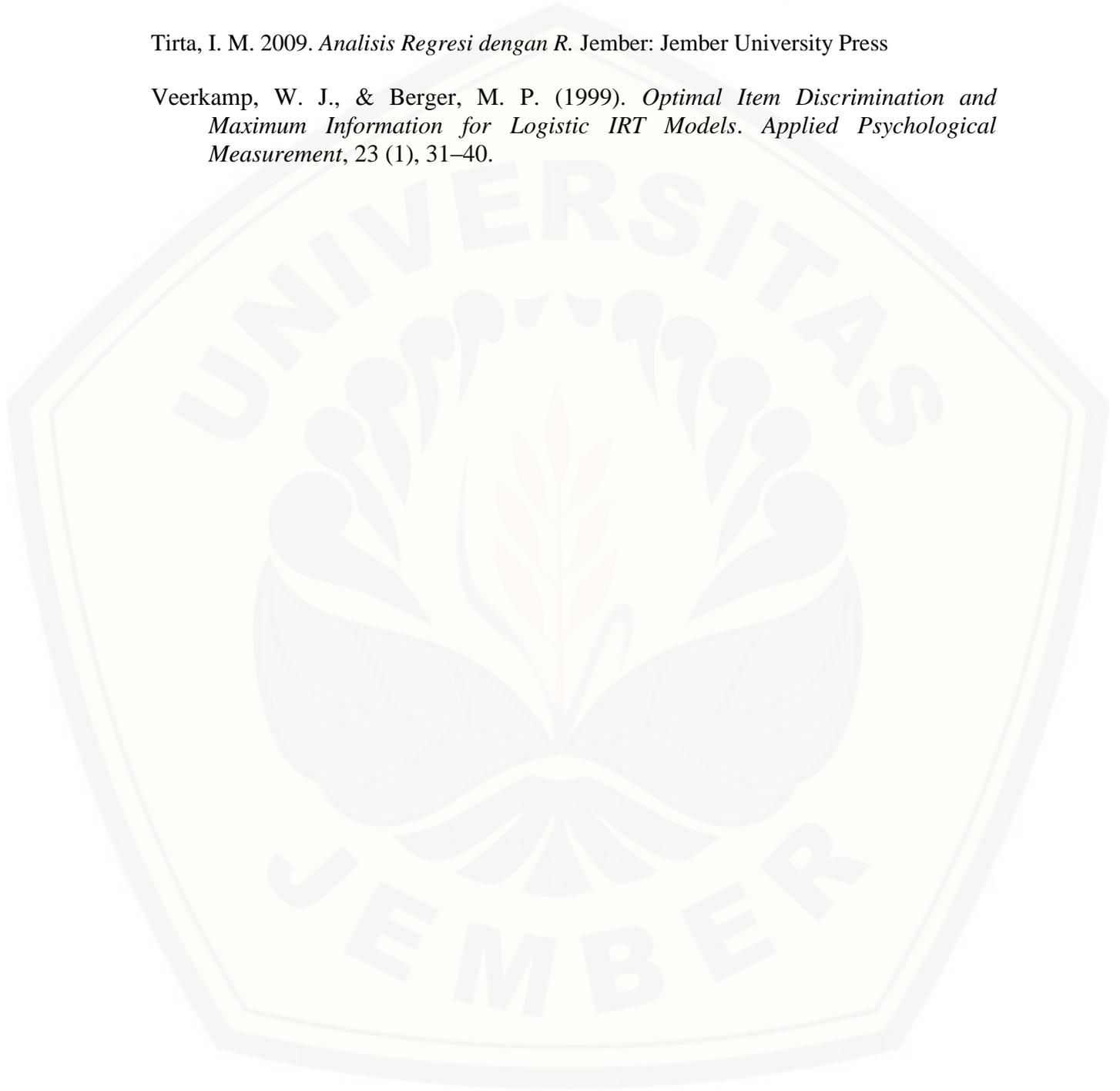
DAFTAR PUSTAKA

- Anastasi, A & S. Urbina. 1997. *Psychological Testing*. Seventh Edition. Prentice Hall International Inc. New Jersey.
- Crocker, L. & J. Algina. 1986. *Introduction to Classical and Modern Test Theory*. Holt, Rinehart and Winston, Inc. New York.
- Embretson, SE & Reise, S.P. 2000. *Item response theory for psychologists*. Artikel. New Jersey: Lawrance Erlbaum Associates.
- Hambleton, R. K., Swaminathan, H. & H. J. Rogers. 1991. *Fundamentals of Item Response Theory*. Sage Publications, The International Profesional Publshers. California.
- Hogg, R. V., McKean, J. W., & A. T. Craig. 1978. *Introduction to Mathematical Statistics 4th Ed*. Macmillan Publishing. New York.
- Isgiyanto, A. 2011. *Analisis Data Ujian Nasional Matematika Berdasarkan Penskoran Model Rasch dan Model Partial Credit*. Skripsi. Yogyakarta: FMIPA Universitas Negeri Yogyakarta.
- Kartono. 2008. *Penyetaraan Tes Model Campuran Butir Dikotomus dan Politomus Pada Tes Prestasi Belajar*. Skripsi. Semarang: FMIPA Universitas Negeri Semarang.
- Lord, F. M. 1980. *Application of Item Response Theory to Practical Testing Problems*. Hillsdale, New Jersey: Lawrence Erlbaum Associates Publishers.
- Nitko, A. J. 1996. *Educational Assesment of Students*. Second Edition. Merrill an imprint of Prentice Hall Englewood Cliffs, Ohio.
- Nugraheni, V. 2010. *Teori Respon Item dengan Pendekatan Model Logistik Satu Parameter*. Skripsi. Semarang: FMIPA Universitas Diponegoro.
- Rizopoulus, D. 2006. *ltm: An R Package for Latent Variable Modeling and Item Response Theory Analyses*. *Journal of Statistical Software*: Volume 17, Issue 5.

Safari, M. 2009. *Teknik Analisis Butir Soal, Instrumen Tes dan Non-Tes dengan Manual, Kalkulator dan Komputer*. Jakarta: Puspendik Depdiknas RI.

Tirta, I. M. 2009. *Analisis Regresi dengan R*. Jember: Jember University Press

Veerkamp, W. J., & Berger, M. P. (1999). *Optimal Item Discrimination and Maximum Information for Logistic IRT Models*. *Applied Psychological Measurement*, 23 (1), 31–40.



3. Indikator : Menyelesaikan masalah yang berkaitan dengan barisan bilangan dan deret

Indikator soal : Menentukan suku ke-n dari barisan aritmatika yang diberikan

Soal :

Dari barisan aritmatika, suku ke-2 = 11 dan suku ke-5 = 23. Suku ke-30 barisan tersebut adalah ...

- a. 111 c. 123
b. 117 d. 135

4. Indikator : Menyelesaikan masalah yang berkaitan dengan peluang suatu kejadian

Indikator soal : Menentukan peluang suatu kejadian tertentu pada suatu percobaan melambungkan tiga uang logam.

Soal :

Dalam percobaan melambungkan 3 uang logam, peluang muncul ketiganya angka adalah...

- a. $\frac{2}{3}$ c. $\frac{3}{8}$
b. $\frac{1}{3}$ d. $\frac{1}{8}$

5. Indikator : Menghitung luas permukaan bangun ruang sisi datar dan sisi lengkung

Indikator soal : Menghitung luas permukaan tabung, kerucut, atau bola

Soal :

Luas permukaan tabung tanpa tutup yang panjang jari-jarinya 7cm dan tingginya 10 cm adalah...

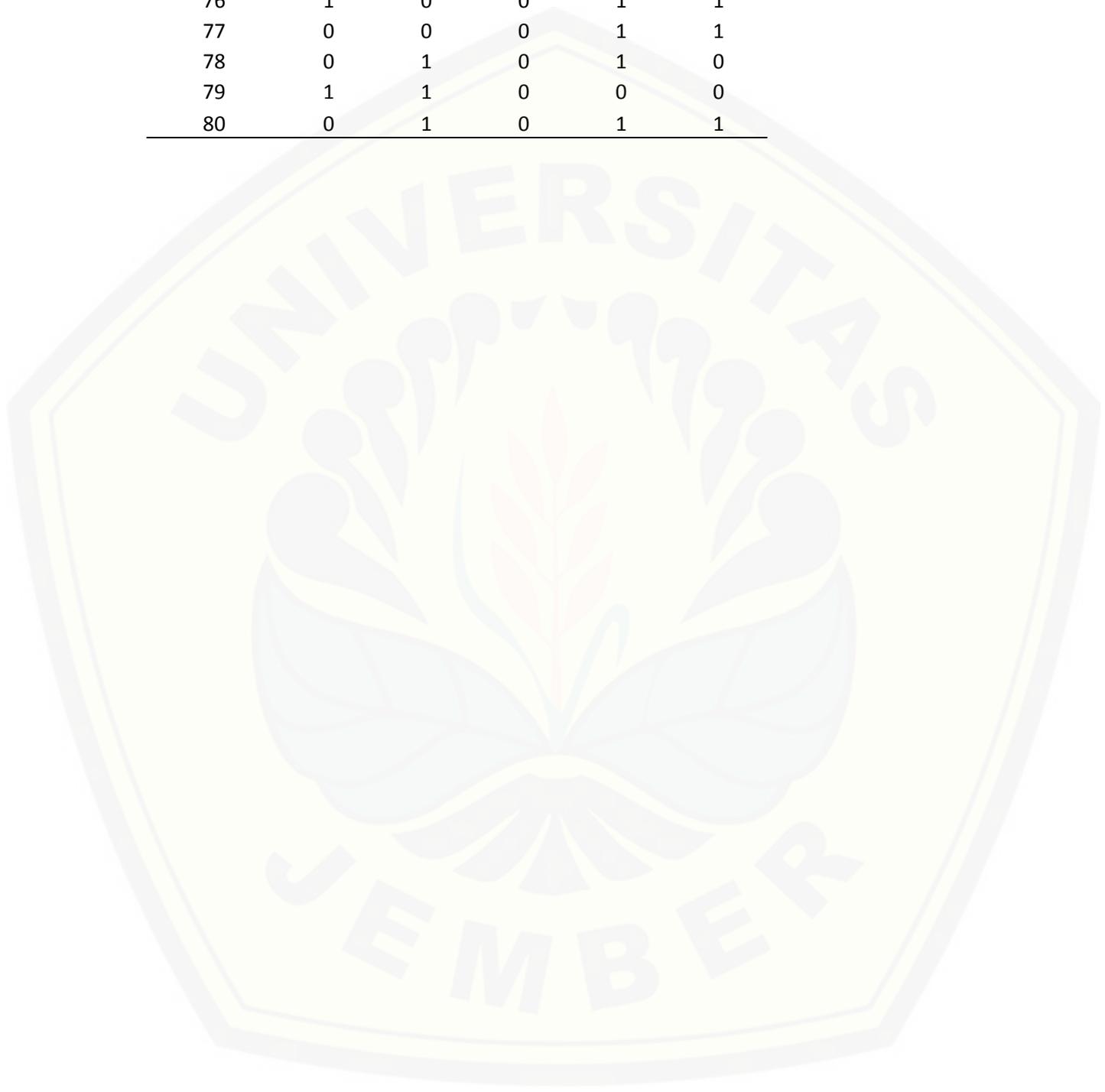
- a. 154 cm^2 b. 440 cm^2 c. 594 cm^2 d. 748 cm^2

LAMPIRAN B. Respon Peserta

Responden	Item 1	Item 2	Item 3	Item 4	Item 5
1	0	1	0	1	1
2	0	1	0	1	0
3	1	1	0	1	1
4	1	1	0	1	1
5	0	1	0	1	1
6	0	1	0	1	1
7	0	1	1	1	0
8	0	1	0	1	1
9	1	1	1	1	1
10	1	1	1	1	1
11	0	1	0	1	1
12	0	0	0	1	0
13	0	1	0	1	0
14	0	0	0	0	1
15	0	1	0	1	1
16	1	1	0	1	1
17	0	0	0	0	0
18	0	1	0	0	0
19	0	1	0	0	0
20	0	0	0	0	0
21	0	0	1	1	0
22	0	1	0	0	0
23	0	1	0	1	0
24	0	1	0	1	0
25	0	1	0	1	0
26	1	1	0	1	0
27	0	0	0	1	0
28	0	0	0	1	0
29	0	0	0	1	0
30	0	0	0	1	0
31	0	0	0	1	0
32	0	1	0	1	0
33	0	1	0	1	0
34	0	1	0	0	0
35	0	0	0	0	0
36	0	0	0	1	0
37	0	1	0	1	0

Responden	Item 1	Item 2	Item 3	Item 4	Item 5
38	0	1	1	1	1
39	0	1	0	1	1
40	0	1	0	1	1
41	1	1	0	1	1
42	0	0	1	1	1
43	0	0	0	1	1
44	0	0	0	1	1
45	1	0	0	1	1
46	0	0	0	1	1
47	1	1	0	1	1
48	1	0	0	1	1
49	1	1	0	1	1
50	1	1	0	1	1
51	1	1	0	0	0
52	0	1	0	1	0
53	0	0	0	1	1
54	0	0	0	1	0
55	0	1	0	1	0
56	0	0	0	0	0
57	0	1	0	1	0
58	0	1	1	1	1
59	0	1	1	0	1
60	0	1	0	0	1
61	0	0	1	0	1
62	0	1	1	1	1
63	0	1	0	1	1
64	0	0	0	1	1
65	0	0	0	1	1
66	0	1	0	1	1
67	0	1	1	1	1
68	0	1	0	1	1
69	0	1	0	1	0
70	0	1	0	1	1
71	0	1	0	1	0
72	0	1	0	0	0
73	0	1	1	1	1
74	0	1	0	1	1
75	1	1	1	1	1

Responden	Item 1	Item 2	Item 3	Item 4	Item 5
76	1	0	0	1	1
77	0	0	0	1	1
78	0	1	0	1	0
79	1	1	0	0	0
80	0	1	0	1	1



LAMPIRAN C. Skrip Program untuk *Rasch* Model Menggunakan *Software R*

```
#####  
### Penerapan Teori Respon Item dengan Rasch Model  
#####  
library(ltm)  
data(datafix)  
  
#####  
### Analysis of datafix  
#####  
  
# descriptive statistics  
descript(datafix)  
  
#####  
# constrained Rasch model #  
#####  
fit1 <- rasch(datafix, constraint = cbind(length(datafix) + 1, 1))  
summary(fit1)  
  
# coefficients and conditional probabilities  
coef(fit1, prob = TRUE, order = TRUE)  
  
# Goodness-of-Fit using Bootstrap  
GoF.rasch(fit1, B = 65)  
  
# Goodness-of-Fit in the margins  
margins(fit1)  
margins(fit1, type = "three-way", nprint = 2)  
  
#####  
# unconstrained Rasch model #  
#####  
fit2 <- rasch(datafix)  
summary(fit2)  
  
# Goodness-of-Fit in the margins  
margins(fit2)  
margins(fit2, type = "three-way", nprint = 2)  
  
# likelihood ratio test between constrained and unconstrained models  
anova(fit1, fit2)  
  
#####  
# Extensions #  
#####  
  
# Two-Parameter Logistic model  
fit3 <- ltm(datafix ~ z1)  
anova(fit1, fit3)
```

```

# Rasch model with guessing parameter
fit4 <- tpm(datafix, type = "rasch", max.guessing = 1)
anova(fit1, fit4)

#####
# Item Characteristic and Information Curves #
#####

# put all plots in one figure
par(mfrow = c(1, 1))

# by default the Item Characteristic Curves are plotted;
# we include a legend
plot(fit1, legend = TRUE, cx = "bottomright", lwd = 3,
     cex.main = 1.5, cex.lab = 1.3, cex = 1.1)

# to plot the Item Information Curves the `type' argument is used;
# since we included a legend in the ICCs plot no extra annotation is
# required --> `annot = FALSE'
plot(fit1, type = "IIC", legend = TRUE, cx = "bottom", annot =
     FALSE, lwd = 3, cex.main = 1.5,
     cex.lab = 1.3)

# the Test Information Curve is plotted by using both `type = "IIC"
# and `items = 0'
plot(fit1, type = "IIC", legend = TRUE, items = 0, lwd = 3, cex.main
     = 1.5,
     cex.lab = 1.3)

# we also include the amount of Test information in the intervals
# (-4, 0) and (0, 4)
plot(0:1, 0:1, type = "n", ann = FALSE, axes = FALSE)
info1 <- information(fit1, c(-4, 0))
info2 <- information(fit1, c(0, 4))
text(0.5, 0.5, labels = paste("Total Information:",
    round(info1$InfoTotal, 3),
    "\n\nInformation in (-4, 0):", round(info1$InfoRange, 3),
    paste("(", round(100 * info1$PropRange, 2), "%)", sep = "")),
    "\n\nInformation in (0, 4):", round(info2$InfoRange, 3),
    paste("(", round(100 * info2$PropRange, 2), "%)", sep = "")),
    cex = 1.5)

```