
Estimasi Tingkat Kesukaran Butir Antara Item Response Theory Dan Rasch Model

Asmin Lukman¹ dan Safri Tinamba²

^{1,2}Universitas Bumi Hijrah Tidore

¹Email. asminlukman@gmail.com

Info Artikel

Kirim: 9 Oktober
2023

Terima: 29 Desember
2023

Terbit Online
Desember 2023

Kata-kata kunci:

Analisis Butir,
Tingkat Kesukaran,
Item Response
Theory, Rasch Model

ABSTRAK

Penelitian ini menggunakan pendekatan deskriptif dengan tujuan untuk menyajikan hasil analisis butir soal satu parameter logistik (1-PL) menggunakan dua pendekatan yaitu pendekatan Item Respons Thoery dan Rasch model. Analisis pendekatan model IRT menggunakan bantuan software BILOG_MG dan Model Rasch menggunakan bantuan software Winsteps. Analaisis satu parameter yang dimaksud adalah tingkat kesukaran butir soal. Subjek dalam penelitian penelititan ini adalah peserta didik yang mengikuti ujian akhir mata pelajaran matematika kelas XII. Jumlah sampel dalam penelitian sebanyak 29 peserta. Pengumpulan data menggunakan pendekatan dokumentasi. Data diperoleh melalui google form yang berisi lembar jawaban dan kunci jawaban. Hasil penelitian mmenunjukkan bahwa terdapat angka tingkat kesukaran butir soal pada model IRT lebih besar dari angka tingkat kesukaran butir pada model rasch. Sebanyak 35 butir soal yang dianalisis memiliki selisih yang berbeda. Selisih yang paling kecil terdapat pada butir soal nomor 3 sebesar 0,01. Indeks kesukaran pada model IRT sebesar 1,27 dan indeks kesukaran pada model Rasch sebesar 1,26. Selisih terbesar terdapat pada butir soal nomor 12 sebesar 1,25. Indeks kesukaran pada model IRT sebesar -0,29 dan indeks kesukaran pada model Rasch sebesar -1,54. Angka tingkat kesukaran butir soal semakin kecil maka selisih tingkat kesukaran IRT dan Rasch makin besar. Sebaliknya, Angka tingkat kesukaran butir soal semakin besar maka selisih tingkat kesukaran IRT dan Rasch makin kecil. Meskipun begitu, IRT dan Rasch memiliki kesamaan apabila butir soal diurutkan mulai dari angka terkecil hingga angka yang terbesar. Kedua model tersebut memiliki urutan soal yang sama yaitu terkecil mulai dari butir soal nomor 12 dan yang terbesar yaitu butir soal nomor 3.

1. PENDAHULUAN

Analisis butir soal menjadi bagian penting dari penilaian tingkat satuan pendidikan. Penilaian yang baik akan menghasilkan informasi yang baik dan dapat

dijadikan sumber pengambilan keputusan. Informasi yang baik sangat ditentukan oleh hasil tes dan analisis butir soal dan analisis butir dapat memastikan bahwa hasil tes benar-benar objektif. Namun pada kenyataannya, kesalahan dalam melakukan penilaian masih seringkali terjadi di satuan pendidikan. Istiyono E, (2014) mengungkapkan bahwa esesmen yang dilakukan sering tidak menginterpretasi kemampuan peserta didik yang sesungguhnya. Istiyono, E dan Ray, M (2020) juga mengungkapkan bahwa para pengajar yang mengabaikan cara penilaian hasil belajar karena masih menggunakan soal-soal yang tidak mencerminkan kemampuan peserta didik. Hal ini dapat terjadi karena para guru masih mengabaikan satu tahapan yang penting dalam pelaksanaan penilaian yaitu analisis butir soal Tujuan menganalisis butir soal yaitu agar dapat melihat indeks reliabilitas tes, validitas setiap butir soal, tingkat kesukaran, daya beda, dan hasil tebakan (guessing). Pendekatan yang dilakukan untuk menganalisis butir soal diantaranya adalah teori tes klasik, rasch model, dan teori respon butir.

Analisis hasil tes dengan teori tes klasik dilakukan dengan mengolah data melalui statistik yang umum digunakan seperti, distribusi frekwensi, proporsi, rata-rata, simpangan baku, variansi, korelasi. Pada teori tes klasik, terdapat formula yang terdiri dari tiga hal pokok yaitu skor amatan R (skor amatan), T (true score), dan E (error). Skor amatan didefinisikan sebagai skor yang diperoleh individu dalam menjawab seluruh butir tes yang dikerjakannya tanpa peduli benar karna paham atau tidak paham. Oleh karena itu, skor amatan belum tentu merepresantasi kemampuan yang sebenarnya. True skor atau skor sebenarnya adalah skor yang diperoleh individu dalam menjawab soal dengan benar karena benar-benar dipahami dan sesuai dengan kemampuannya. Error atau kesalahan adalah skor benar yang diperoleh seseorang namun benar tersebut bukan karena kemampuannya. Kesalahan skor ini sulit untuk diketahui karena hasilnya selalu berubah-ubah.

Teori Tes Klasik Memiliki dua kelemahan (Mardapi, J., 2012). Kelemahan dari teori tes klasik ini adalah tes sangat tergantung pada kemampuan kelompok dan tingkat kesukaran butir. Artinya bahwa keberhasilan peserta tes sangat bergantung pada kelompok butir dalam perangkat tes. Tingkat kesukaran butir juga sangat bergantung pada kelompok responden. Dengan kata lain, apabila tes itu diberikan kepada siswa

yang memiliki kemampuan tinggi maka butir menjadi mudah, dan apabila tes diberikan pada siswa yang memiliki kemampuan rendah maka butir menjadi sukar. Kemudian butir mudah diberikan pada peserta tes diberikan ke siswa yang kemampuannya rendah maka skor yang diperoleh tinggi, sebaliknya butir sukar diberikan kepada siswa yang kemampuannya tinggi maka skor yang diperoleh juga tinggi. Sehingga dapat disimpulkan bahwa kedua kelompok tersebut memiliki kemampuan yang sama dan tidak dapat dibedakan.

Analisis butir pada teori respon butir berfungsi untuk mendeteksi butir-butir soal yang baik untuk dijadikan perangkat tes. Analisis dilakukan untuk menganalisis dan mengkaji tiap butir tes dengan bantuan statistik pada responden untuk tiap butir soal. Kualitas butir soal dapat dilihat pada nilai informasi butir soal. Informasi butir pada teori respon butir menunjukkan hubungan antara parameter butir dengan ability pada tes. parameter butir berupa tingkat kesukaran, daya beda dan tebakan. Indeks tingkat kesukaran (b) butir berada pada $-2 \leq b \leq 2$. Indeks daya beda (a) berada pada rentang $0 \leq a \leq 2$, dan indeks tebakan (c) berada diantara $0 \leq c \leq 0,5$. Parameter ability (θ) berada pada rentang $-4 \leq \theta \leq 4$.

Kualitas suatu perangkat tes sangat tergantung pada kualitas tiap butir soal. kualitas butir tes untuk model logistik memiliki dua parameter, yaitu parameter butir dan parameter ability. Parameter butir bergantung pada model logistik yang digunakan. Model parameter logistik terdiri dari 1 PL, 2 PL, dan 3 PL. Rumus probabilitas menjawab benar model logistik

- Rumus satu parameter logistik

$$P_j(\theta) = \frac{1}{1 + e^{-D(\theta - b_j)}}$$

- Rumus dua parameter logistik

$$P_j(\theta) = \frac{1}{1 + e^{-D a_j (\theta - b_j)}}$$

- Rumus tiga parameter logistik

$$P_j(\theta) = c_j + (1 - c_j) \frac{1}{1 + e^{-D a_j (\theta - b_j)}}$$

P_j = probabilitas peserta tes ke-j dengan taraf kemampuan (θ_j) menjawab butir ke-j dengan benar; a_j = daya beda butir ke-j; b_j =tingkat kesukaran butir ke-j; c_j = menjawab benar dengan tebakan butir ke-j; e = besaran konstanta 2,718; D = besaran konstanta 1,702

- Pemodelan Rasch (1-PL) data dikotomi

Model pengukuran Rasch (1-PL) pertamakali dikembangkan oleh Rasch (Rasch, 1980). Data yang berbentuk dikotomi, pemodelan rasch menurut Bambang & Wahyu (2015:36) menggabungkan satu algoritma yang menyatakan hasil ekspektasi probabilistik dari butir ‘I’ dan responden ‘n’, yang secara matematis dinyatakan sebagai:

$$P_{ni}(X_n=1 | \beta_n, \delta_i) = \frac{e^{(b_n - d_i)}}{1 + e^{(b_n - d_i)}}$$

dimana : $P_{ni}(X_n=1 | b_n, d_i)$ adalah probabilitas dari responden n dalam butir i untuk menghasilkan jawaban betul ($x=1$): dengan kemampuan responden $=\beta_n$, dan tingkat kesulitan butir $= \delta_i$. Persamaan di atas oleh Rasch dapat lebih disederhanakan dengan memasukkan fungsi logaritma dan menjaidikannya :

$$\log (P_{ni}(X_{ni} = 1 | \beta_n, \delta_i))= \beta_n - \delta_i$$

Menurut Bambang dan Wahyu (2015: 46) keunggulan pemodelan Rasch dibandingkan pemodelan lain yaitu; a) Kemampuan memprediksi data yang hilang yang didasarkan pada pola respon yang sistematis. b) Pemodelan rasch mampu menghasilkan kemungkinan nilai terbaik dari data yang hilang. c) Kaliberasi dalam Pemodelan rasch secara sekaligus dalam tiga hal yaitu, skala pengukuran, responden, dan butir soal.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan penelitian deskriptif kualitatif. penelitian ini bertujuan untuk mengetahui hasil analisis butir 1-PL menggunakan dua pendekatan yaitu pendekatan Item Respons Thoery (IRT) dan Rasch model . Subjek dalam penelitian penelititan ini adalah peserta didik yang mengikuti ujian akhir mata pelajaran matematika. Jumlah sampel dalam penelitian sebanyak 29 peserta.

Pengumpulan data menggunakan pendekatan dokumentasi. Data diperoleh melalui google form yang berisi lembar jawaban siswa, dilengkapi dengan kunci jawabannya.

Analisis butir menggunakan bantuan dua software analisis butir yaitu software BILOG MG untuk analisis butir model item reponse theory dan software winsteps untuk analisis pada model rash. Kedua software ini memiliki kesamaan yaitu dapat menghasilkan tingkat kesukaran butir soal. Hasil analisis kedua software ini kemudian dilakukan perbandingan yaitu antara tingkat kesukaran butir pada Software BILOG MG dan tingkat kesukaran butir model rasch pada software Winsteps. Jenis data yang dianalisis adalah data dikotomi karena hanya ada jawaban yaitu benar dan salah.

Setelah menganalisis dengan menggunakan kedua software tersebut maka selanjutnya akan diambil hasil dari indeks kesukaran tiap butir soal antara IRT (BILOG_MG) dan Rasch (Winstep). Setelah itu dilihat selisih angka antara IRT dan Rasch mulai dari butir soal nomor 1 sampai butir soal nomor 35. Tingkat kesukaran butir pada software Winstep akan dilihat pada Item Measura (Salsabila, F, 2023). Dan tingkat kesukaran butir pada software BILOG_MG akan dilihat pada kolom logit.

3. HASIL DAN PEMBAHASAN

Analisis data diawali dengan mempersiapkan data terlebih dahulu yakni dibuat dalam file bentuk excell. Data yang disalin di excel berupa jawaban peserta didik dalam bentuk huruf A, B, C, D, atau E. Kemudian file tersebut disalin kembali dalam bentuk notepad lalu disiapkan untuk dianalisis. Berikut ini adalah hasil response (jawaban) peserta didik yang sudah dibuat dalam bentuk file notepad.

NOMOR01BCCCBBCAEDBDABDACDBACDBBCADAB
NOMOR02EDDACEABAEB CABADDDECABBODDACC
NOMOR03EDDOEEEDDBECABAAAACAADDEBBEBCB
NOMOR04DDDBDBBBBDCBBDDDDDEBAABEDCBCB
NOMOR05AAADABBDAACBCADAOCACEA0BABAAD
NOMOR06AEEDBDBBEABDCBECDDDEDECBDDBAEAD
NOMOR07CACCAECEEBBEADBCEDDBCACCADBEC
NOMOR08EDDEEEDDDBEDABAAAACAADDEBBECCD
NOMOR09EBECABCCBCCBCCDCCDDABEBBDCCCC
NOMOR10DCCDCBCCCEDEACEAAEBDBDCCDECAA
NOMOR11EADBADDCCBDBEDBBCEBBADCBACAAAD
NOMOR12BABBBBCBDBCCBCCDCCDBCDBDBCBBD
NOMOR13ADDBEBBBBBDCDEEABCBDDBABAEACEC
NOMOR14CEEDDABCEABCDCDDADBBBCBCBCBBAD
NOMOR15AAABOBCAACCABCABCDAABAABABBABB
NOMOR16CEEEEAABCCBCEDCBBAACBDABEECBB
NOMOR17CDDDCBCBCCCCCCCCCCCCDCCBCCC
NOMOR18BDBDDBEBCADBBBDEDDBECAEDCAADD
NOMOR19ADDAADBACEBBEADACDBBEBBDAEDAC
NOMOR20CEEDDBBBCDBCCCBABBACEABCBBECC
NOMOR21CBBDCEBEDACCBBCCBDBCAABODABDAB
NOMOR22CEEDDADEBOCCACEEBADODACBOBCAA
NOMOR23CEEDDACCEABCDBDAEADACDBBADAB
NOMOR24AEEDBAEDCAACBCDCAAEDCBADABDD
NOMOR25CEEDDACDCEBCADEBABDBDADEEADAE
NOMOR26AEEDDEAAAABCBEBAABCDBAACBDDDEE
NOMOR27CEEDDACBCCBCDBBEDBBACBECCBBED
NOMOR28CEEDEAEBDBBCEBCBCDAACDADBBBAD
NOMOR29CCEDDBEDECDCBBBAEABDDDEABCBED

Gambar 1 di atas menunjukkan jumlah peserta didik yang mengikuti tes sebanyak 29 orang. Kode peserta didik yaitu mulai dari NOMOR01 artinya orang pertama sampai orang ke 29 yang disimbolkan dengan NOMOR29. Kemudian jumlah soal yang dikerjakan dalam bentuk pilihan ganda sebanyak 35 butir soal. Setiap butir dapat dijawab oleh peserta didik karena soal yang dibuat melalui google form sehingga setiap butir diseting untuk wajib menjawab. Peserta didik tidak dapat menyelesaikan apabila masih ada butir soal yang belum terjawab. Meskipun hanya ada satu butir soal yang tidak terjawab maka hasil jawaban mereka dinyatakan belum selesai.

Langkah selanjutnya menganalisis hasil jawaban peserta didik tersebut menggunakan bantuan software BILOG MG. Analisis butir dengan menggunakan software BILOG MG dapat menghasilkan 3 parameter logistik sekaligus yakni 1 PL berkaitan dengan tingkat kesukaran, 2 PL berkaitan dengan daya beda, dan 3 PL yang berkaitan dengan tebakan. Namun, hasil analisis dengan software BILOG MG ini yang dilihat hanya pada 1PL karena akan dibandingkan dengan 1PL pada rasch model dengan bantuan software winsteps. Berikut ini hasil analisis 1-PL pada model analisis item response theory dengan bantuan software BILOG MG.

Gambar 2. Hasil analisis dengan software Bilog MG

ITEM STATISTICS FOR SUBTEST TEST0001

ITEM	NAME	#TRIED	#RIGHT	PCT	LOGIT/1.7	ITEM*TEST CORRELATION	
						PEARSON	BISERIAL
1	ITEM0001	29.0	12.0	41.4	0.20	-0.042	-0.053
2	ITEM0002	29.0	11.0	37.9	0.29	0.173	0.220
3	ITEM0003	29.0	3.0	10.3	1.27	0.157	0.265
4	ITEM0004	29.0	7.0	24.1	0.67	0.298	0.409
5	ITEM0005	29.0	12.0	41.4	0.20	0.312	0.395
6	ITEM0006	29.0	13.0	44.8	0.12	0.052	0.065
7	ITEM0007	29.0	8.0	27.6	0.57	0.200	0.267
8	ITEM0008	29.0	9.0	31.0	0.47	0.364	0.477
9	ITEM0009	29.0	6.0	20.7	0.79	0.204	0.290
10	ITEM0010	29.0	9.0	31.0	0.47	0.027	0.035
11	ITEM0011	29.0	7.0	24.1	0.67	0.200	0.275
12	ITEM0012	29.0	18.0	62.1	-0.29	0.343	0.438
13	ITEM0013	29.0	9.0	31.0	0.47	0.248	0.325
14	ITEM0014	29.0	4.0	13.8	1.08	-0.254	-0.398
15	ITEM0015	29.0	8.0	27.6	0.57	-0.301	-0.403
16	ITEM0016	29.0	5.0	17.2	0.92	0.268	0.397
17	ITEM0017	29.0	7.0	24.1	0.67	0.399	0.547
18	ITEM0018	29.0	3.0	10.3	1.27	0.058	0.098
19	ITEM0019	29.0	5.0	17.2	0.92	0.026	0.039
20	ITEM0020	29.0	7.0	24.1	0.67	-0.235	-0.323
21	ITEM0021	29.0	9.0	31.0	0.47	0.203	0.265
22	ITEM0022	29.0	10.0	34.5	0.38	-0.006	-0.007
23	ITEM0023	29.0	5.0	17.2	0.92	0.052	0.078
24	ITEM0024	29.0	5.0	17.2	0.92	0.105	0.156
25	ITEM0025	29.0	6.0	20.7	0.79	-0.115	-0.163
26	ITEM0026	29.0	12.0	41.4	0.20	0.401	0.507
27	ITEM0027	29.0	8.0	27.6	0.57	-0.026	-0.034
28	ITEM0028	29.0	7.0	24.1	0.67	-0.012	-0.016
29	ITEM0029	29.0	11.0	37.9	0.29	0.325	0.415
30	ITEM0030	29.0	6.0	20.7	0.79	0.230	0.326
31	ITEM0031	29.0	10.0	34.5	0.38	-0.148	-0.191
32	ITEM0032	29.0	11.0	37.9	0.29	0.580	0.739
33	ITEM0033	29.0	14.0	48.3	0.04	-0.052	-0.065
34	ITEM0034	29.0	9.0	31.0	0.47	-0.121	-0.158
35	ITEM0035	29.0	9.0	31.0	0.47	-0.121	-0.158

Hasil analisis di atas menunjukkan bahwa susunan butir pada kolom NAME dimulai dari soal nomor 1 dan terakhir nomor 35 artinya bahwa butir disusun berdasarkan urutan nomor soal. Tingkat kesukaran butir terdapat pada kolom LOGIT. Nomor butir yang paling sulit adalah nomor 18 dan 3 karena indeks kesukaran butir sebesar 1,27. Butir soal yang paling sulit adalah nomor 12 karena indeks kesukaran butirnya sebesar -0,29.

Gambar 3. Hasil Analisis Butir dengan Software Winsteps

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S. E.
3	3	29	1.26	.62
18	3	29	1.26	.62
14	4	29	.92	.55
16	5	29	.65	.50
19	5	29	.65	.50
23	5	29	.65	.50
24	5	29	.65	.50
9	6	29	.41	.47
25	6	29	.41	.47
30	6	29	.41	.47
4	7	29	.20	.44
11	7	29	.20	.44
17	7	29	.20	.44
20	7	29	.20	.44
28	7	29	.20	.44
7	8	29	.01	.43
15	8	29	.01	.43
27	8	29	.01	.43
8	9	29	-.16	.41
10	9	29	-.16	.41
13	9	29	-.16	.41
21	9	29	-.16	.41
34	9	29	-.16	.41
35	9	29	-.16	.41
22	10	29	-.33	.40
31	10	29	-.33	.40
2	11	29	-.49	.40
29	11	29	-.49	.40
32	11	29	-.49	.40
1	12	29	-.64	.39
5	12	29	-.64	.39
26	12	29	-.64	.39
6	13	29	-.79	.39
33	14	29	-.94	.38
12	18	29	-1.54	.40
MEAN	8.4	29.0	.00	.44
S. D.	3.2	.0	.60	.06

Gambar di atas menunjukkan bahwa pada kolom ENTRY NUMBER berisi ‘nomor’ soal. Pada entri number menampilkan soal dimulai dari nomor 3 sampai nomor 12 yang artinya bahwa susunan nomor soal tidak berurutan. Hasil analisis dengan winstep menampilkan ‘tingkat kesulitan’ soal pada kolom MEASURE butir soal diurutkan mulai dari soal yang paling sulit hingga soal yang paling mudah. Soal yang paling sulit adalah nomor 3 karena hanya mampu dijawab dengan benar oleh 3 siswa. Soal yang paling mudah yaitu soal nomor 12 karena mampu dijawab dengan benar oleh 18 siswa.

Perbandingan hasil analisis butir dengan menggunakan software bilog dan winsteps akan dilihat indeks kesukaran butir. Indeks kesukaran butir dengan model Item response theory pada 1PL dan indeks kesukaran butir model rasch 1-PL. Agar lebih jelas perbandingan dua model analisis butir tersebut dapat dibuat dalam bentuk tabel sebagai berikut.

Tabel 1. Tingkat kesukaran pada IRT dan Rasch model

Nomor Soal	Indek kesukaran (IRT)	Indek Kesukaran (Rasch)	Selisih
1	0,20	-0,64	0,84
2	0,29	-0,49	0,78
3	1,27	1,26	0,01
4	0,67	0,20	0,47
5	0,20	-0,64	0,84
6	0,12	-0,79	0,91
7	0,57	0,01	0,56
8	0,47	-0,16	0,63
9	0,79	0,41	0,38
10	0,47	-0,16	0,63
11	0,67	0,20	0,47
12	-0,29	-1,54	1,25
13	0,47	-0,16	0,63
14	1,08	0,92	0,16
15	0,57	0,01	0,56
16	0,92	0,65	0,27
17	0,67	0,20	0,47
18	1,27	1,26	0,01
19	0,92	0,65	0,27
20	0,67	0,20	0,47
21	0,47	-0,16	0,63
22	0,38	-0,33	0,71
23	0,92	0,65	0,27
24	0,92	0,65	0,27
25	0,79	0,41	0,38
26	0,20	-0,64	0,84
27	0,57	0,01	0,56
28	0,67	0,20	0,47
29	0,29	-0,49	0,78
30	0,79	0,41	0,38
31	0,38	-0,33	0,71
32	0,29	-0,49	0,78
33	0,04	-0,94	0,98
34	0,47	-0,16	0,63
35	0,47	-0,16	0,63

Indeks kesukaran tiap butir pada IRT dan Rasch model berbeda angkanya, seperti pada butir soal nomor 1, indeks kesukaran pada IRT memiliki angka 0,20 dan pada rasch model memiliki indeks kesukaran sebesar -0,64. Artinya bahwa selisih dari kedua indeks kesukaran butir soal tersebut sebesar 0,84. Kriteria tingkat kesukaran pada rasch (Bambang & Swamintono, 2015) adalah ($b > 0,80$) sangat sukar, ($0 \leq b \leq 0,8$) sukar, ($-0,8 \leq b < 0$) mudah, dan ($b < -0,80$) sangat mudah. Kriteria tingkat kesukaran pada IRT adalah butir dikatakan baik apabila memiliki indeks kesukaran diantara -2 dan +2 ($-2 \leq b \leq 2$).

Selisih yang paling kecil terdapat pada butir soal nomor 3 karena memiliki selisih sebesar 0,01. Kemudian butir soal yang memiliki selisih paling besar yaitu butir soal nomor 12 dengan selisih sebesar 1,27. Dari selisih tersebut dapat dinyatakan bahwa semakin sulit butir soal maka selisih tingkat kesukaran kedua model tersebut semakin kecil. Sebaliknya semakin mudah butir soal maka selisih tingkat kesukaran dari kedua model tersebut makin besar. Berdasarkan probabilitas 1-PL pada model IRT dan Rasch model juga sedikit perbedaan pada rumus probabilitasnya yaitu pada IRT terdapat nilai $-D=-1,702$ sementara pada Rasch model tidak terdapat nilai $-D=-1,702$.

$$P_j(\theta) = \frac{1}{1 + e^{-D(\theta - b_j)}} \quad (\text{Model IRT})$$

$$P_{ni}(X_n=1 \mid \beta_n, \delta_i) = \frac{e^{(b_n - d_i)}}{1 + e^{(b_n - d_i)}} \quad (\text{Rasch Model})$$

Rumus pertama dan rumus kedua di atas terkait probabilitas kemampuan individu menjawab benar pada tiap butir soal. Probabilitas kemampuan individu menjawab benar pada butir soal sangat tergantung pada parameter kemampuan dan parameter kesukaran butir baik model IRT maupun model rasch. Indeks kesukaran butir semakin tinggi maka selisihnya hampir sama, akan tetapi jika indeks kesukaran semakin rendah maka selisihnya semakin besar. Jadi tingkat kesukaran sudah sesuai dengan kedua rumus probabilitas tersebut.

4. KESIMPULAN

Hasil analisis butir pada model IRT dengan bantuan software BILOG_MG dan model Rasch dengan bantuan software Winsteps menunjukkan perbedaan angka pada indeks kesukaran butir. Butir 1 sampai butir ke 35 tidak ada yang sama angka pada indeks kesukaran. Artinya, setiap butir memiliki selisi angka indeks kesukaran yang berbeda antara Irt dan Rasch. Semakin kecil indeks kesukaran pada butir soal tertentu maka selisihnya semakin besar, atau sebaiknya semakin besar angka tingkat kesukaran butir maka selisihnya semakin kecil. Meskipun begitu, apabila butir soal diurutkan mulai dari angka terkecil hingga angka yang terbesar maka kedua model tersebut memiliki urutan yang sama. Hal ini terlihat bahwa kedua model tersebut sama-sama memperoleh tingkat kesukaran paling tinggi yaitu butir nomor 3 dan tingkat kesukaran paling rendah terdapat pada butir soal nomor 12.

DAFTAR PUSTAKA

- Fadiya S., Juntika, dan Yaya, S (2023). Pengujian Validitas dan Reliabilitas Instrumen Menejmen diri Remaja : Rasch Model Analysis. *Jurnal Bimbingan dan Konseling terapan* Volume 07, nomor 01, 15-25
- Istiyono, E., & Ray M. (2020). Analisis Penilaian Soal Fisika Menggunakan Model Rasch dengan Program R. *Jurnal Sains dan Edukasi Sains. Vol 3, no.2 Agustus 2020.* 46-52.
- Istiyono, E. (2014). *Kemampuan Berpikir Tingkat Tinggi Fisika Peserta Didik SMA di DIY*. Yogyakarta: Desertasi UNY.
- Mardapi, D (2012). *Pengukuran, Penilaian & Evaluasi Pendidikan*. Yogyakarta : Nuha Medika
- Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. Chicago : University of Chicago Press.
- Sumintono, B., & Widiharso, B. (2015). *Aplikasi Pemodelan Rasch pada Asesmen Pendidikan*. Cimahi: Trim Komunikata.