

**STUDI KOMPARATIF: SENSITIVITAS UNTUK MENDETEKSI
BIAS BUTIR PADA METODE UBTS, KHI-KUADRAT LORD DAN
DISTRIBUSI SAMPLING EMPIRIS**



**I Wayan Widana
NIM. 7817130252**

Program Studi: Penelitian dan Evaluasi Pendidikan

Disertasi yang Ditulis untuk Memenuhi Sebagian Persyaratan
untuk Mendapatkan Gelar Doktor

**PROGRAM PASCASARJANA
UNIVERSITAS NEGERI JAKARTA
2016**

**STUDI KOMPARATIF: SENSITIVITAS UNTUK MENDETEKSI BIAS BUTIR
PADA METODE UJI BEDA TARAF SUKAR, KHI-KUADRAT LORD DAN
DISTRIBUSI SAMPLING EMPIRIS
(2016)**

**COMPARISON STUDY: SENSITIVITY OF DETECTION ITEM BIAS OF
TEST OF DIFFICULTY DIFFERENCE METHOD, LORD CHI-SQUARE, AND
EMPIRICAL SAMPLING DISTRIBUTIONS
(2016)**

I WAYAN WIDANA

ABSTRACT

The objective of this research was to understand the sensitivity difference of DIF methods: (1) test of b difference, (2) Lord's chi-square, and (3) empirical sampling distributions to detect item bias. This research is a descriptive exploratory.

Data were obtained from Puspendik. It was response data of a high school student science major in DKI Jakarta Province amount 4,869 people. The item bias test being tested was the national exam of mathematical, academic year 2014/2015 package code 1101. To detect item bias, replication is done through 30, the sample randomly selected of both the reference group and focal group each 1000 people. Data were analyzed using the Bilog-MG Program, SPSS 22.0, and Microsoft Excel 2010. The hypothesis testing using One_Way ANOVA statistical and Tukey Post Hoc test.

The results showed that: (1) the empirical sampling distributions method is more sensitive than the Lord's chi-square method to detect item bias with mean difference 3.200 and value-sig.0,000<0.05; (2) the Lord's chi-square method more sensitive than the test of b difference method for detecting item bias, with mean difference of 2.367 and value-sig.0,001<0.05.

Keyword: *DIF Methods test of b difference, Lord's chi-square, empirical sampling distributions, item bias.*

RINGKASAN

PENDAHULUAN

Hasil pengukuran diharapkan dapat memberikan informasi yang sesuai dengan karakteristik peserta tes tanpa merugikan atau menguntungkan individu atau kelompok tertentu akibat ketidakadilan alat ukur tersebut. Perangkat tes dikatakan memuat bias butir, jika memuat butir-butir yang memihak kelompok tertentu. Bias butir diartikan sebagai ketidakvalidan suatu butir, sehingga butir tersebut tidak memberikan peluang yang sama menjawab benar pada dua kelompok yang berbeda pada peserta dengan kemampuan yang sama.

Pendeteksian bias butir pada penelitian ini menggunakan 3 metode *DIF* berdasarkan pendekatan *IRT* model L3P, sebagaimana direkomendasikan oleh Camilli dan Shepard (1994): UBTS, Khi-Kuadrat Lord, dan DSE. Sebelum proses mendeteksi bias butir dilakukan, terlebih dahulu parameter butir dan kemampuan dari dua kelompok peserta tes yang berbeda, skalanya disamakan. Metode penyetaraan yang digunakan untuk penyamaan skala adalah metode Rerata dan Sigma. Pada metode UBTS bias butir dideteksi menggunakan rumus: $d = \frac{\Delta \hat{b}}{S_{\Delta \hat{b}}}$. Jika nilai $d > z = 1,96$ atau

$d < z = -1,96$ pada taraf signifikansi $\alpha = 0,05$, maka butir ke-*i* dinyatakan memuat bias butir. Pada metode Khi-Kuadrat Lord, bias butir dideteksi menggunakan rumus: $\chi_i^2 = \mathbf{v}_i' \Sigma_i^{-1} \mathbf{v}_i$. Jika $\chi_{hitung}^2 > 7,81472$ dengan $dk=3$, $\alpha = 0,05$, maka butir ke-*i* dinyatakan memuat bias butir. Pada metode DSE, terlebih dahulu dihitung indek bias butir antarkelompok R-F (δ) menggunakan rumus: $SPD - \theta = \frac{\sum_{j=1}^{n_p} \Delta P(\theta_j)}{n_p}$ untuk bias butir *uniform* dan $UPD - \theta =$

$\sqrt{\frac{\sum_{j=1}^{n_p} (\Delta P(\theta_j))^2}{n_p}}$ untuk bias butir *non-uniform*, dengan: $\Delta P = P_R(\theta_j) - P_F(\theta_j)$.

Selanjutnya dihitung indeks bias butir subkelompok R1-R2 (δ_1), dan F1-F2 (δ_2). Jika nilai $\delta >$ nilai maksimum ($\delta_1; \delta_2$), maka butir ke-i dinyatakan memuat bias butir.

METODE PENELITIAN

Penelitian ini merupakan penelitian deskriptif eksploratif, yaitu penelitian yang mendeskripsikan kinerja metode *DIF* dan perbandingan sensitivitasnya untuk mendeteksi bias butir. Populasi berjumlah 4.869 orang siswa SMA Program IPA Provinsi DKI Jakarta, yang terdiri dari 2.114 orang siswa laki-laki (kelompok referensi) dan 2.755 orang siswa perempuan (kelompok fokal). Tes yang digunakan adalah soal UN Matematika SMA IPA tahun pelajaran 2014/2015 kode paket 1101.

Untuk menguji sensitivitas masing-masing metode *DIF* dilakukan 30 replikasi, jumlah sampel kelompok referensi dan fokal masing-masing 1000 orang yang diambil secara acak. Data dianalisis menggunakan Program BILOG-MG, SPSS 22.0, dan Microsoft Excel 2010. Pengujian hipotesis menggunakan statistik ***One-Way ANOVA*** dan uji ***Post Hoc*** menggunakan Uji ***Tukey***.

HASIL PENELITIAN

Data hasil pendeteksian bias butir masing-masing metode *DIF*, dianalisis melalui dua tahap menggunakan Program SPSS 22.0 ***One_Way ANOVA***. Hasil analisis tahap pertama:

Tabel ANOVA

BIAS_BUTIR					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	468.289	2	234.144	38.452	.000
Within Groups	529.767	87	6.089		
Total	998.056	89			

Pada Tabel Anova di atas, nilai F sebesar 38,452 dengan signifikansi sig.=0,000<0,05. Kesimpulan: tolak H_0 dan terima H_1 , artinya terdapat perbedaan yang signifikan rerata jumlah bias butir yang dideteksi menggunakan metode UBTS, Khi-Kuadrat Lord, dan DSE. Dengan demikian Uji **Post Hoc** dapat dilanjutkan, dengan hasil sebagai berikut.

Tabel Multiple Comparisons

Dependent Variable: BIAS_BUTIR

Tukey HSD

(I) METODE	(J) METODE	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
UBTS	LORD	-2.367*	.637	.001	-3.89	-.85
	DSE	-5.567*	.637	.000	-7.09	-4.05
LORD	UBTS	2.367*	.637	.001	.85	3.89
	DSE	-3.200*	.637	.000	-4.72	-1.68
DSE	UBTS	5.567*	.637	.000	4.05	7.09
	LORD	3.200*	.637	.000	1.68	4.72

*. The mean difference is significant at the 0.05 level.

Pengujian Hipotesis

1. Metode DSE lebih sensitif daripada metode Khi-Kuadrat Lord untuk mendeteksi bias butir. Pada Tabel **Multiple Comparisons** terlihat bahwa metode DSE dan metode Khi-Kuadrat Lord memiliki perbedaan rerata 3,200 dengan nilai sig.0,000<0,05, kesimpulan: tolak H_0 dan terima H_1 . **Hipotesis pertama teruji.**
2. Metode Khi-Kuadrat Lord lebih sensitif daripada metode UBTS untuk mendeteksi bias butir. Pada Tabel **Multiple Comparisons** terlihat bahwa Khi-Kuadrat Lord dan metode UBTS memiliki perbedaan rerata sebesar 2,367 dengan nilai sig.0,001<0,05 (signifikan), kesimpulan: tolak H_0 dan terima H_1 . **Hipotesis kedua teruji.**

Kesimpulan dan Implikasi:

1. Metode DSE lebih sensitif daripada metode Khi-Kuadrat Lord untuk mendeteksi bias butir.
2. Metode Khi-Kuadrat Lord lebih sensitif daripada metode UBTS untuk mendeteksi bias butir.
3. Bagi para guru dan praktisi pendidikan direkomendasikan untuk melakukan pendeteksian bias butir menggunakan metode DSE, karena lebih sensitif dibandingkan dengan metode Khi-Kuadrat Lord dan UBTS.
4. Untuk memperkuat hasil penelitian ini, para peneliti diharapkan dapat mengembangkan penelitian sejenis dengan jumlah replikasi yang lebih besar.

Lembar Persetujuan Ujian Tertutup Disertasi

PERSETUJUAN KOMISI PROMOTOR DIPERSYARATKAN UNTUK UJIAN TERTUTUP DISERTASI	
Promotor	Co-Promotor
Prof. Dr. H. Djaali 2016	Prof. Dr. Ir. Dali S. Naga, MMSI. 2016
Mengetahui, Ketua Program Studi PEP PPs Universitas Negeri Jakarta Prof. Dr. Burhanuddin Tola, M.A. 2016	
Nama	: I Wayan Widana
No. Registrasi	: 7817130252
Tanggal Lulus	:

LEMBAR PERNYATAAN

Saya menyatakan dengan sesungguhnya bahwa Disertasi yang saya susun sebagai syarat untuk memperoleh gelar Doktor dari Program Pascasarjana Universitas Negeri Jakarta seluruhnya merupakan hasil karya saya sendiri.

Adapun bagian-bagian tertentu dalam penulisan Disertasi yang saya kutip dari hasil karya orang lain telah dituliskan sumbernya secara jelas sesuai norma, kaidah, dan etika penulisan ilmiah.

Apabila dikemudian hari ditemukan seluruh atau sebagian Disertasi ini bukan hasil karya saya sendiri atau adanya plagiat dalam bagian-bagian tertentu, saya bersedia menerima sanksi pencabutan gelar akademik yang saya sandang dan sanksi-sanksi lainnya sesuai dengan peraturan perundangan yang berlaku.

Jakarta, April 2016

I Wayan Widana

KATA PENGANTAR

Rasa syukur serta segala puja dan puji dipanjatkan dihadapan Tuhan Yang Maha Esa, karena berkat rahmat-Nya Disertasi yang berjudul “Studi Komparatif: Sensitivitas untuk Mendeteksi Bias Butir pada Metode Uji Beda Taraf Sukar, Khi-Kuadrat Lord dan Distribusi Sampling Empiris” dapat diselesaikan sesuai dengan rencana. Disertasi ini ditulis untuk memenuhi sebagian persyaratan untuk mendapatkan gelar Doktor dari Program Pascasarjana Universitas Negeri Jakarta.

Disertasi ini dapat diselesaikan atas bantuan dari berbagai pihak. Pada kesempatan yang baik ini, ucapan terima kasih dan penghargaan yang setinggi-tingginya disampaikan kepada:

1. Prof. Dr. H. Djaali, selaku Rektor Universitas Negeri Jakarta sekaligus sebagai promotor yang telah banyak memberikan bimbingan dan arahan.
2. Prof. Dr. Ir. Dali S. Naga, MMSI., selaku co-promotor yang telah banyak meluangkan waktu untuk berdiskusi, memberikan petunjuk, dan bimbingan.
3. Prof. Dr. Burhanuddin Tola, M.A., selaku Kaprodi PEP Program Pascasarjana UNJ yang telah berkenan memberikan petunjuk dan arahan.

4. Dr. Wardani Rahayu, M.Si., selaku Sekretaris Prodi PEP Program Pascasarjana UNJ yang telah berkenan meluangkan waktu untuk berdiskusi, memberikan petunjuk, dan arahan.
5. Dr. Heri Retnawati, M.Pd., dosen Program Pascasarjana UNY atas kesabaran dan keikhlasannya telah banyak memberikan kontribusi untuk memperdalam penguasaan materi psikometrika dan penggunaan program (*software*) aplikasi pengukuran.
6. Kepala Pusat Penilaian Pendidikan (Puspendik) Balitbang Kemendikbud, atas dukungannya berupa penyediaan data UN Mata Pelajaran Matematika Program IPA Provinsi DKI Jakarta Tahun pelajaran 2014/2015.
7. Luh Srinadi istri tercinta, anak-anakku tersayang I Putu Wahyu Budhi Sucita dan Adek Widya P. Saraswati, serta kedua orang tua dan seluruh keluarga yang telah memberikan dukungan moral, doa, dan materi.

Disertasi ini masih sangat jauh dari sempurna, oleh karena itu masukan dan saran konstruktif dari berbagai pihak sangat diharapkan untuk penyempurnaan penulisan-penulisan di masa yang akan datang.

Jakarta, April 2016

I Wayan Widana

DAFTAR ISI

	Hal.
ABSTRAK	ii
RINGKASAN.....	iii
PERSETUJUAN KOMISI PROMOTOR	vii
PERSETUJUAN PERBAIKAN	viii
LEMBAR PERNYATAAN	ix
KATA PENGANTAR	x
DAFTAR ISI	xii
DAFTAR TABEL	xvi
DAFTAR GAMBAR/BAGAN.....	xvii
DAFTAR LAMPIRAN	xix
BAB I PENDAHULUAN.....	1
A. Latar Belakang.....	1
B. Identifikasi Masalah	12
C. Pembatasan Masalah	14
D. Rumusan Masalah.....	15
E. Kegunaan Penelitian.....	15

BAB II KAJIAN TEORETIK	18
A. Deskripsi Konseptual	19
1. Teori Responsi Butir (<i>Item Response Theory/IRT</i>)	19
a. Konsep Dasar <i>IRT</i>	19
b. Asumsi-asumsi <i>IRT</i>	23
c. Parameter Responden dan Butir	28
d. Estimasi Parameter	30
e. Model Logistik	36
f. Pencocokan Model	45
g. Fungsi Informasi	48
2. Metode Penyetaraan (<i>Linking Method</i>)	51
3. Bias Butir	58
a. Pengertian	59
b. Jenis-jenis Bias Butir	66
c. Bias Gender	68
d. Metode <i>DIF</i> yang Diperbandingkan	74
1) Metode Uji Beda Taraf Sukar	77
2) Metode Khi-Kuadrat Lord	79

3) Metode Distribusi Sampling Empiris	81
e. Perbandingan Sensitivitas Metode <i>DIF</i>	85
B. Hasil Penelitian yang Relevan	85
C. Kerangka Teoretik	91
D. Hipotesis Penelitian	99
BAB III METODOLOGI PENELITIAN.....	100
A. Tujuan Penelitian	100
B. Tempat dan Waktu Penelitian	100
C. Metode Penelitian	101
D. Populasi dan Sampel.....	102
E. Teknik Pengumpulan Data.....	103
F. Teknik Analisis Data.....	108
G. Hipotesis Statistik	116
BAB IV HASIL PENELITIAN DAN PEMBAHASAN.....	118
A. Deskripsi Data	118
B. Pengujian Persyaratan Analisis	136
C. Pengujian Hipotesis.....	138
D. Pembahasan Hasil Penelitian.....	142

BAB V KESIMPULAN, IMPLIKASI, DAN SARAN	179
A. Kesimpulan	179
B. Implikasi.....	180
C. Saran.....	181
DAFTAR PUSTAKA.....	182

DAFTAR TABEL

	Hal.
1. Tabel 2.1 Ringkasan Studi terdahulu	86
2. Tabel 3.1 Rancangan Penelitian	102
3. Tabel 3.2 Tabel Bantu Anava Satu Jalur	114
4. Tabel 4.1 Data Peserta UN SMA-IPA Paket 1101	119
5. Tabel 4.2 Indikator pada Kisi-Kisi UN 2014/2015.....	120
6. Tabel 4.3 Uji Kecukupan Sampel.....	123
7. Tabel 4.4 Pencocokan Model.....	124
8. Tabel 4.5 Kode Butir yang Cocok Model.....	126
9. Tabel 4.6 Parameter Butir Soal UN Paket 1101.....	127
10. Tabel 4.7 Kategori Taraf Suksar	129
11. Tabel 4.8 Nilai Fungsi Informasi Butir Maksimum	130
12. Tabel 4.9 Bias Butir yang Terdeteksi	132
13. Tabel 4.10 Ringkasan Banyak Bias Butir pada Metode <i>DIF</i>	134
14. Tabel 4.11 Uji Homogenitas.....	137
15. Tabel 4.12 Uji Perbedaan Rerata.....	139
16. Tabel 4.13 Hasil Uji <i>Post Hoc</i>	141

DAFTAR GAMBAR/BAGAN

	Hal.
1. Gambar 2.1 Distribusi Kondisional Skor Tes	26
2. Gambar 2.2 Peta Alir Langkah Iterasi	35
3. Gambar 2.3 Kurva Karakteristik Butir Model L1P	40
4. Gambar 2.4 Kurva Karakteristik Butir Model L2P	42
5. Gambar 2.5 Kurva Karakteristik Butir Model L3P	44
6. Gambar 2.6 Butir Tidak Bias	65
7. Gambar 2.7 Butir Bias	65
8. Gambar 2.8 Bias Butir <i>Uniform</i>	67
9. Gambar 2.9 Bias Butir <i>Non-Uniform</i>	68
10. Gambar 4.1 Kurva karakteristik Butir Kode 19	147
11. Gambar 4.2 Kurva Karakteristik Butir Kode 14	153
12. Gambar 4.3 Kurva Karakteristik Butir Kode 1	157
13. Gambar 4.4 Grafik Fungsi Informasi Butir Kode 1	158
14. Gambar 4.5 Kurva Karakteristik Butir Kode 13	162
15. Gambar 4.6 Grafik Fungsi Informasi Butir Kode 13	163
16. Gambar 4.7 Kurva Karakteristik Butir Kode 24	167

17. Gambar 4.8 Grafik Fungsi Informasi Butir Kode 24	168
18. Gambar 4.9 Kurva karakteristik Butir Kode 23	174

DAFTAR LAMPIRAN

	Hal.
1. Lampiran 1. Uji Kecocokan Model dan Parameter Butir	188
2. Lampiran 2. Nilai Fungsi Informasi Butir	191
3. Lampiran 3. Syntax BILOG-MG	194
4. Lampiran 4. Hasil Replikasi.....	195
5. Lampiran 5. Surat Permohonan Izin Penelitian	
6. Lampiran 6. Berita Acara Serah Terima Hasil UN	
7. Lampiran 7. Soal UN Matematika SMA IPA Kode 1101	
8. Lampiran 8. Daftar Riwayat Hidup	

BAB I

PENDAHULUAN

A. LATAR BELAKANG

Evaluasi pendidikan adalah kegiatan pengendalian, penjaminan, dan penetapan mutu pendidikan terhadap berbagai komponen pendidikan pada setiap jalur, jenjang, dan jenis pendidikan sebagai bentuk pertanggungjawaban penyelenggaraan pendidikan. Oleh karena itu, perlu ditetapkan standar penilaian pendidikan yang terkait dengan kriteria mengenai mekanisme, prosedur, dan instrumen penilaian hasil belajar peserta didik.¹

Keberhasilan suatu pendidikan dapat diukur melalui kegiatan evaluasi pendidikan. Evaluasi pendidikan merupakan bagian yang sangat penting dan tidak terpisahkan dalam proses pembelajaran. Dengan melakukan evaluasi, guru sebagai pengelola kegiatan pembelajaran dapat mengetahui kemampuan yang dimiliki peserta didik, ketepatan metode mengajar yang digunakan, dan keberhasilan peserta didik untuk mencapai kompetensi yang telah ditetapkan. Berdasarkan hasil evaluasi, guru dapat mengambil keputusan secara tepat untuk menentukan langkah yang harus dilakukan

¹ *Peraturan Pemerintah Republik Indonesia Nomor 13 Tahun 2015 tentang Perubahan Kedua Atas Peraturan Pemerintah Republik Indonesia Nomor 19 Tahun 2005 tentang Standar Nasional Pendidikan (SNP), pasal 1.*

selanjutnya. Hasil evaluasi juga dapat memberikan motivasi kepada peserta didik untuk berprestasi lebih baik.²

Penilaian hasil belajar peserta didik mencakup kompetensi sikap, pengetahuan, dan keterampilan yang dilakukan secara berimbang sehingga dapat digunakan untuk menentukan posisi relatif setiap peserta didik terhadap standar yang telah ditetapkan. Penilaian pendidikan sebagai proses pengumpulan dan pengolahan informasi untuk mengukur pencapaian hasil belajar peserta didik mencakup: penilaian autentik, penilaian diri, penilaian berbasis portofolio, ulangan, ulangan harian, ulangan tengah semester, ulangan akhir semester, ujian tingkat kompetensi, ujian mutu tingkat kompetensi, ujian nasional, dan ujian sekolah/madrasah.³

Untuk menjamin agar penilaian dapat memberikan informasi yang benar, maka penilaian hendaknya dilakukan menggunakan prinsip-prinsip penilaian antara lain: (1) sahih, berarti penilaian didasarkan pada data yang mencerminkan kemampuan yang diukur; (2) objektif, berarti penilaian didasarkan pada prosedur dan kriteria yang jelas, tidak dipengaruhi subjektivitas penilai; (3) adil, berarti penilaian tidak menguntungkan atau merugikan peserta didik karena berkebutuhan khusus serta perbedaan latar belakang agama, suku, budaya, adat istiadat, status sosial ekonomi, dan

² Tim Penyusun, *Rancangan Penilaian* (Jakarta: Direktorat Pembinaan SMA, 2010), h. 1.

³ *Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 66 Tahun 2014 tentang Standar Penilaian Pendidikan*, hh. 2-3.

gender; (4) terpadu, berarti penilaian oleh pendidik merupakan salah satu komponen yang tak terpisahkan dari kegiatan pembelajaran; (5) terbuka, berarti prosedur penilaian, kriteria penilaian, dan dasar pengambilan keputusan dapat diketahui oleh pihak yang berkepentingan; (6) menyeluruh dan berkesinambungan, berarti penilaian oleh pendidik mencakup semua aspek kompetensi dan dengan menggunakan berbagai teknik penilaian yang sesuai dengan kompetensi yang harus dikuasai peserta didik; (7) sistematis, berarti penilaian dilakukan secara berencana dan bertahap dengan mengikuti langkah-langkah baku; (8) beracuan kriteria, berarti penilaian didasarkan pada ukuran pencapaian kompetensi yang ditetapkan; dan (9) akuntabel, berarti penilaian dapat dipertanggungjawabkan baik dari segi teknik, prosedur, maupun hasilnya.⁴

Sesuai dengan prinsip-prinsip penilaian di atas, Pusat Penilaian Pendidikan Balitbang Kementerian Pendidikan dan Kebudayaan (2013) menyatakan bahwa terdapat dua permasalahan pokok yang harus dipertimbangkan untuk mendapatkan hasil pengukuran yang baik yaitu: (1) alat ukur atau tes yang digunakan menghasilkan skor yang reliabel, dan (2) alat ukur tersebut harus valid, artinya dapat mengukur dengan tepat sesuai

⁴ *Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 53 Tahun 2015 tentang Penilaian Hasil Belajar oleh Pendidik pada Pendidikan Dasar dan Pendidikan Menengah, pasal 4, hh. 4-5.*

dengan target objek (kemampuan peserta tes) yang akan diukur.⁵ Apabila alat ukur telah memenuhi kedua kriteria tersebut, akan diperoleh hasil pengukuran ulang yang sesuai dengan tujuan pengukuran tanpa terpengaruh oleh faktor-faktor lainnya. Hasil pengukuran diharapkan dapat memberikan informasi yang sesuai dengan karakteristik peserta tes tanpa merugikan atau menguntungkan individu atau kelompok-kelompok tertentu akibat ketidakadilan alat ukur tersebut.

Alat ukur yang baik hendaknya dapat memberikan informasi secara akurat terkait dengan atribut yang diukur dalam suatu pengukuran. Dalam teori responsi butir, informasi ini dikenal sebagai nilai fungsi informasi. Fungsi informasi butir (*item information function*) merupakan suatu metode untuk menjelaskan kekuatan atau kontribusi suatu butir pada perangkat tes dalam mengungkap *latent trait* yang diukur menggunakan tes tersebut (fungsi informasi tes).⁶ Nilai fungsi informasi tes merupakan penjumlahan nilai fungsi informasi butir. Semakin tinggi nilai fungsi informasi masing-masing butir pada perangkat tes, maka semakin tinggi pula nilai fungsi informasi tes, dan sebaliknya. Dengan demikian, pencapaian hasil belajar peserta didik dapat dijelaskan oleh nilai fungsi informasi tes. Artinya, sejauh mana tes yang digunakan untuk mengukur hasil belajar dapat menjelaskan kemampuan

⁵ Puspendik Balitbang Kemdikbud, *Panduan Pengembangan dan Pemberdayaan Bank Soal di Daerah* (Jakarta: Puspendik, 2013), h. 8.

⁶ Heri Retnawati, *Teori Responsi Butir dan Penerapannya Untuk Peneliti, Praktisi Pengukuran dan Pengujian, Mahasiswa Pascasarjana* (Yogyakarta: Nuha Medika, 2014), h. 18.

peserta didik. Pencapaian hasil belajar peserta didik dinyatakan dalam bentuk kompetensi dasar. Sehingga apabila suatu tes disusun oleh butir-butir soal yang memiliki fungsi informasi yang tinggi, maka tes tersebut diharapkan dapat menjelaskan sejauh mana kompetensi dasar yang seharusnya dicapai oleh peserta didik sudah tercapai.

Ujian Nasional yang selanjutnya disebut UN adalah kegiatan pengukuran dan penilaian pencapaian standar kompetensi lulusan secara nasional pada mata pelajaran tertentu.⁷ Salah satu mata pelajaran yang diujikan dalam UN adalah Matematika. Hal itu didasarkan atas pertimbangan bahwa mata pelajaran matematika merupakan ilmu universal yang mendasari perkembangan teknologi modern, serta mempunyai peranan penting dalam berbagai disiplin ilmu dan memajukan daya pikir manusia. Untuk menguasai dan menciptakan teknologi di masa depan, diperlukan penguasaan dan pemahaman atas matematika yang kuat sejak dini. Matematika dapat membangun kreativitas dan inovasi peserta didik, sehingga diharapkan dapat meningkatkan daya saing dengan bangsa-bangsa lain di dunia.⁸

Mengingat sedemikian pentingnya mata pelajaran matematika untuk membangun logika, berpikir kritis, dan kreatif agar memiliki daya saing yang

⁷ *Lampiran Peraturan Badan Standar Nasional Pendidikan Nomor 0031/P/BSNP/III/2015 tentang Prosedur Operasional Standar Penyelenggaraan Ujian Nasional Tahun Pelajaran 2014/2015*, h. 2.

⁸ *Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 59 Tahun 2014 tentang Kurikulum 2013 Sekolah Menengah Atas/Madrasah Aliyah, Lampiran III*, h. 325.

tinggi di tingkat internasional, maka pengukuran terhadap pencapaian hasil belajar matematika hendaknya dilakukan dengan objektif dan adil (tidak menguntungkan atau merugikan kelompok-kelompok tertentu), sehingga hasil pengukuran dapat menggambarkan pencapaian kompetensi peserta didik. Untuk dapat mengungkap kompetensi matematika yang telah dimiliki peserta didik, perlu diukur menggunakan instrumen yang memenuhi standar mutu di antaranya memiliki nilai fungsi informasi yang tinggi serta memenuhi kaidah-kaidah penulisan butir soal yang baik.

Di dalam kenyataan, ada kalanya skor hasil tes itu tidak memberikan informasi yang benar terhadap karakteristik peserta tes. Mungkin saja informasi tersebut tidak menjangkau sampai ke besaran atau dimensi yang hendak diukur oleh tes tersebut. Mungkin pula hasil tes tersebut tercampur dengan besaran atau dimensi lain yang tidak dimaksudkan untuk diukur oleh tes tersebut sehingga hasil tes menjadi rancu. Akibatnya skor yang diperoleh tidak benar atau terjadi ketimpangan skor. Ketimpangan skor tidak memberikan informasi yang benar tentang hal-hal yang dimaksud untuk diukur oleh uji tes itu.⁹ Dalam penilaian, ketimpangan skor hendaknya dapat diminimalkan agar pengukuran dapat memberikan gambaran yang utuh tentang karakteristik objek yang diukur.

⁹ Dali S. Naga, *Pengantar Teori Sekor pada Pengukuran Pendidikan* (Jakarta: Besbats, 1992), hh. 435-436.

Ada banyak sebab yang menjurus ke ketimpangan skor dalam pengukuran. Namun, secara garis besar dikenal dua sumber atau wilayah tempat ketimpangan skor itu terjadi yaitu peserta tes dan perangkat tes beserta butir-butirnya. Selain sumber atau tempat kejadian ketimpangan skor, juga dapat dilihat bentangan dari ketimpangan skor itu. Pada peserta tes, ketimpangan dapat terjadi secara individu atau pada kelompok. Demikian pula tes, ketimpangan dapat terjadi pada butir atau keseluruhan tes. Dengan demikian terdapat empat kombinasi ketimpangan skor: (1) ketimpangan skor pada peserta secara individu, (2) ketimpangan skor pada peserta secara kelompok, (3) ketimpangan skor pada butir secara individu satu demi satu, dan (4) ketimpangan skor pada butir secara kelompok atau tes. Selanjutnya ketimpangan skor pada kombinasi pertama disebut ketidakwajaran, kombinasi ketiga disebut bias butir, dan kombinasi keempat disebut bias tes.¹⁰

Jika pada suatu tes memuat butir-butir yang memihak kelompok tertentu, maka tes tersebut dikatakan memuat bias butir atau mengandung keberfungsian butir diferensial (*Differential Item Functioning, DIF*). Bias butir hendaknya diminimalkan atau dihilangkan sama sekali. Terutama pada tes yang digunakan secara luas dan hasilnya sangat menentukan masa depan seseorang seperti Ujian Sekolah (US), UN, SBMPTN, dan berbagai bentuk

¹⁰ *Ibid*, hh. 436-437.

tes seleksi lainnya. Tes yang digunakan harus bebas dari bias butir yang dapat merugikan individu atau sekelompok individu tertentu. Sebelum tes beserta butir-butirnya digunakan, seharusnya diujicobakan terlebih dahulu sehingga apabila terdapat butir-butir yang mengandung bias dapat terdeteksi sejak awal.

Berbagai upaya telah dilakukan untuk meminimalkan adanya bias butir. Studi yang dilakukan oleh Tess Miller, Saad Chahine & Ruth A. Childs (2010) mengilustrasikan penggunaan *DIF* dan *Differential Step Functioning (DSF)* untuk mendeteksi perbedaan taraf sukar butir yang digunakan pada ujian Matematika siswa kelas IX di Ontario.¹¹ Demikian pula Marie Wiberg (2009) melakukan penelitian tentang perbandingan tiga metode *DIF* untuk mendeteksi bias butir.¹² Youngsuk Suh & Daniel M. Bolt (2011), melakukan penelitian tentang pengaruh pengecoh (*disctractor*) terhadap bias butir pada soal pilihan ganda.¹³ Hasil penelitiannya menyatakan bahwa pengecoh berpengaruh terhadap bias butir.

¹¹ Tess Miller, Saad Chahine & Ruth A. Childs, Detecting Differential Item Functioning and Differential Step Functioning Due to Differences that Should Matter. *Journal Practical Assessment, Research & Evaluation, Volume 15, Number 10, July 2010, ISSN 1531-7714*, h. 2. (diakses 25 Oktober 2015).

¹² Marie Wiberg, Differential Item Functioning in Mastery Tests: A Comparison of Three Methods Using Real Data. *International Journal of Testing, 9: 41–59, 2009*, Taylor & Francis Group, LLC, ISSN: 1530-5058, h. 41, (diakses 9 Oktober 2015).

¹³ Youngsuk Suh & Daniel M. Bolt, A Nested Logit Approach for Investigating Distractors as Causes of Differential Item Functioning. *Journal of Educational Measurement Summer 2011, Vol. 48, No. 2*, h. 188, (diakses 9 Oktober 2015).

Untuk mendeteksi adanya bias butir dalam seperangkat tes, terdapat sejumlah metode yang dapat digunakan. Dalam teori klasik dikenal beberapa metode antara lain: (1) kelompok tunggal (*single group validity*), (2) korelasi diferensial (*differential validity*), (3) model regresi atau model Clearly, (4) prosedur diskriminasi butir (*item discrimination procedure*), (5) metode *plot delta* (*delta plot method*), (6) pendekatan khi-kuadrat Scheuneman (*Scheuneman chi-squared approach*), (7) pendekatan khi-kuadrat Camilli (*Camilli chi-squared approach*). Berbeda dengan skor klasik, skor modern memiliki sifat invarian. Skor peserta invarian terhadap ubahan pada butir tes, serta skor butir invarian terhadap ubahan peserta tes. Secara garis besar ada dua cara yang paling banyak digunakan untuk pendeteksian *DIF*, yaitu: (1) pencocokan parameter ciri butir di antara subpopulasi 1 dan subpopulasi 2 sambil memeriksa kecocokan model; (2) penghitungan luas wilayah di antara lengkungan karakteristik butir yang dibentuk oleh subpopulasi 1 dan subpopulasi 2. Butir tidak bias jika luas itu sama dengan nol atau sangat kecil.¹⁴

Adam E. Wyse & Raymond Mapuranga (2009) merumuskan metode *DIF* baru menggunakan indeks kesamaan informasi (*Information Similarity Index, ISI*). Metode tersebut membandingkan fungsi informasi butir dari data yang cocok dengan model Rasch. Setelah membandingkan metode *ISI*, *Signed Area Index (SAI)*, dan prosedur Mantel-Haenszel untuk mendeteksi

¹⁴ Dali S. Naga, *op. cit.*, h. 472.

bias butir, dinyatakan bahwa metode *ISI* lebih baik daripada metode *SAI* dan prosedur Mantel-Haenszel.¹⁵

Pui-Wa Lei, Shu-Ying Chen, dan Lan Yu (2006) dalam penelitiannya menyatakan bahwa metode *DIF Logistic Regression (LR)* dan *Item Response Theory Likelihood Ratio Test (IRT-LRT)*, lebih sensitif dibandingkan dengan metode *SIBTEST* untuk mendeteksi bias butir tipe tidak konsisten (*non-unidirectional*) pada model *Computerized Adaptive Testing (CAT)*.¹⁶ Oleh karena itu, pada model pengujian adaptif dianjurkan menggunakan metode *LR* dan *IRT-LRT* untuk mendeteksi bias butir.

Sementara itu, Camilli dan Shepard (1994) mengemukakan bahwa terdapat 5 metode untuk mendeteksi bias butir dengan pendekatan teori responsi butir, yaitu: (1) *test of b difference* (uji beda taraf sukar), (2) *item drift methods*, (3) *Lord's chi-square*, (4) *empirical sampling distributions* (distribusi sampling empiris), dan (5) *model comparison measures*.¹⁷ Permasalahan yang muncul kemudian adalah di antara metode-metode pendeteksi *DIF* tersebut di atas, metode manakah yang paling akurat digunakan untuk mendeteksi adanya bias butir?

¹⁵ Adam E. Wyse & Raymond Mapuranga, Differential Item Functioning Analysis Using Rasch Item Information Functions. *International Journal of Testing*, 9: 333–357, 2009, Taylor & Francis Group, LLC, ISSN: 1530-5058 (diakses 9 Oktober 2015), h. 333.

¹⁶ Pui-Wa Lei, Shu-Ying Chen, dan Lan Yu, Comparing Methods of Assessing Differential Item Functioning in a Computerized Adaptive Testing Environment. *Journal of Educational Measurement Vol. 43, No. 3, 2006* (diakses 9 Oktober 2015), h. 245.

¹⁷ Camilli, G. dan Shepard, L. A., *Methods for Identifying Biased Test Item* (California: Sage Publications Inc., 1994), hh. 70-74.

Untuk mendeteksi butir-butir yang mengandung bias butir, sebuah populasi dibagi menjadi dua kelompok yaitu kelompok fokal dan kelompok referensi. Kelompok fokal merupakan kelompok yang akan diselidiki apakah terdapat bias butir pada kelompok itu. Sedangkan kelompok referensi merupakan kelompok pembanding. Kedua kelompok mengerjakan butir pada perangkat tes yang sama. Sebelum proses mendeteksi bias butir dilakukan, terlebih dahulu parameter butir dan parameter responden dari dua kelompok peserta tes yang berbeda skalanya disamakan. Tujuannya adalah agar parameter butir dari dua kelompok peserta tes yang berbeda itu berada pada metrik yang sama, sehingga layak dibandingkan. Metode penyamaan skala itu disebut metode *linking*. Kim dan Lee (2004) mengemukakan bahwa, pada *IRT* terdapat 4 metode *linking* yang sering digunakan oleh para ahli, yaitu: (1) metode rerata dan rerata, (2) rerata dan sigma, (3) *Haebara*, dan (4) *Stocking Lords*.¹⁸

Penyamaan skala (*linking*) dapat dilakukan secara terpisah antara dua kelompok peserta tes atau dapat dilakukan secara simultan tergantung dari metode *DIF* yang digunakan. Misalnya, pada metode *DIF* tes rasio kebolehjadian (*likelihood ratio test*), penyamaan skala dan proses mendeteksi *DIF* dilakukan secara serentak (*concurrent*). Sedangkan pada metode *DIF test of difficulty difference* (uji beda taraf sukar), *item drift*

¹⁸ Kim, S., dan Lee, W. C., *IRT Scale Linking Methods for Mixed-Format Tests* (Iowa: ACT Inc., 2004), hh. iv.

methods, *Lord's chi-square*, dan *empirical sampling distributions for DIF indices* (distribusi sampling empiris), penyamaan skala dan proses mendeteksi *DIF* dilakukan secara terpisah (*sparate*). Artinya, sebelum dilakukan proses mendeteksi bias butir, terlebih dahulu dilakukan penyamaan skala secara terpisah. Masing-masing metode *DIF* yang telah diuraikan di atas, memiliki karakteristik yang khas terkait dengan langkah-langkah pendeteksian bias butir. Dari sekian banyak metode pendeteksi *DIF* yang dapat dipilih, permasalahan yang muncul adalah di antara metode-metode *DIF* yang telah diuraikan di atas, metode manakah yang paling akurat untuk mendeteksi bias butir?

Untuk menjawab permasalahan-permasalahan yang terkait dengan sensitivitas metode pendeteksi *DIF* yang paling akurat digunakan untuk mendeteksi bias butir, perlu dilakukan penelitian. Melalui hasil penelitian dapat dijelaskan secara ilmiah akurasi metode *DIF* tertentu untuk mendeteksi butir-butir yang bias (*invalid*) sehingga dapat dijadikan alternatif oleh *stakeholder* dalam pemilihan metode *DIF*.

B. IDENTIFIKASI MASALAH

Berdasarkan uraian di atas, muncul beberapa permasalahan. Permasalahan-permasalahan itu muncul terkait dengan pemilihan metode-metode *DIF* yang akan digunakan. Bila dilihat dari karakteristiknya masing-

masing metode tersebut memiliki teknik yang berbeda dalam perhitungannya.

Metode *DIF* manakah yang paling sensitif untuk mendeteksi bias butir?

Untuk mendeteksi adanya butir-butir yang mengandung bias butir, dapat dilakukan menggunakan metode-metode *DIF* dengan pendekatan teori tes klasik dan teori responsi butir. Pendekatan teori responsi butir, berusaha meningkatkan akurasi pengukuran melalui pemisahan parameter butir dan parameter kemampuan peserta tes. Hubungan antara parameter butir dan parameter kemampuan dinyatakan dalam bentuk fungsi probabilitas menjawab benar yang dapat digambarkan sebagai kurva karakteristik butir. Permasalahan yang muncul adalah sebagai berikut: (1) apakah metode-metode *DIF* pendeteksi bias butir menggunakan pendekatan teori responsi butir lebih akurat bila dibandingkan dengan metode-metode *DIF* pendeteksi bias butir menggunakan pendekatan teori tes klasik?; (2) apa saja persyaratan yang harus dipenuhi oleh butir agar dapat dideteksi menggunakan metode-metode *DIF* dengan pendekatan teori responsi butir?; (3) apa saja keunggulan-keunggulan dan kelemahan-kelemahan metode-metode *DIF* dengan pendekatan teori tes klasik?; (4) apa saja keunggulan-keunggulan dan kelemahan-kelemahan metode-metode *DIF* dengan pendekatan teori responsi butir? (5) di antara metode-metode *DIF* dengan pendekatan teori responsi butir, metode manakah yang paling akurat untuk mendeteksi bias butir?

C. PEMBATASAN MASALAH

Sesuai dengan identifikasi masalah yang telah dipaparkan di atas, terdapat sejumlah masalah. Namun mengingat keterbatasan-keterbatasan yang ada, maka dalam penelitian ini digunakan pendekatan teori responsi butir dengan model logistik tiga parameter (L3P). Model logistik L3P memiliki tiga parameter yaitu daya pembeda (a), taraf sukar butir (b), dan kebetulan menjawab benar (c).

Di antara metode *DIF* yang menggunakan pendekatan teori responsi butir yang telah dipaparkan di atas, maka yang digunakan untuk mendeteksi bias butir dalam penelitian ini adalah metode Uji Beda Taraf Sukar, Khi-Kuadrat Lord, dan Distribusi Sampling Empiris. Pemilihan tersebut didasarkan pada rekomendasi oleh Camilli dan Shepard (1994), yang menyatakan bahwa metode-metode *DIF* yang berdasarkan teori responsi butir lebih akurat dibandingkan dengan metode *DIF* yang didasarkan pada teori tes klasik. Di samping itu ketiga metode *DIF* Uji Beda Taraf Sukar, Khi-Kuadrat Lord, dan Distribusi Sampling Empiris penyamaan skala dan proses mendeteksi *DIF* dilakukan secara terpisah (*sparate*). Dengan demikian, penelitian ini membandingkan sensitivitas metode *DIF* menggunakan pendekatan teori responsi butir.

D. RUMUSAN MASALAH

Berdasarkan uraian yang telah dipaparkan pada latar belakang, identifikasi masalah, dan pembatasan masalah di atas, maka rumusan masalah dalam penelitian adalah sebagai berikut.

1. Apakah metode Distribusi Sampling Empiris lebih sensitif untuk mendeteksi bias butir daripada metode Khi-Kuadrat Lord?
2. Apakah metode Khi-Kuadrat Lord lebih sensitif untuk mendeteksi bias butir daripada metode Uji Beda Taraf Sukar?

E. KEGUNAAN PENELITIAN

Secara umum penelitian ini diharapkan dapat memberikan dua manfaat, yaitu: (1) manfaat teoretis, dan (2) manfaat praktis. Kedua manfaat tersebut dapat diuraikan sebagai berikut.

1. Manfaat Teoretis

Temuan-temuan dalam penelitian ini diharapkan dapat memberikan kontribusi positif terhadap upaya-upaya untuk pengembangan ilmu pengukuran secara luas. Pengembangan ilmu pengukuran itu sangat penting dilakukan karena banyak digunakan dalam berbagai disiplin ilmu. Kegiatan pengukuran diperlukan dalam pelaksanaan evaluasi terhadap program tertentu. Di samping itu, hasil penelitian ini juga diharapkan dapat memperkuat teori-teori yang telah dikembangkan sebelumnya.

2. Manfaat Praktis

Secara praktis, hasil penelitian ini dapat memberikan informasi yang dapat dimanfaatkan oleh para penyusun tes, pelaksana tes, dan pemangku kebijakan (*stakeholder*) yang terkait dengan pengukuran hasil belajar peserta didik. Informasi-informasi tersebut diharapkan dapat meningkatkan mutu hasil pengukuran. Terlebih lagi hasil pengukuran tersebut akan digunakan secara luas, yang dapat memberikan dampak psikologis dan akademis. Adapun pihak-pihak yang diharapkan dapat memanfaatkan hasil-hasil penelitian ini adalah sebagai berikut.

a. Guru

Guru sebagai ujung tombak dalam pelaksanaan penilaian pendidikan, diharapkan memiliki wawasan dan pemahaman yang memadai untuk melaksanakan pengukuran hasil belajar peserta didik. Dengan wawasan dan pemahaman yang dimiliki tersebut, guru dapat mengembangkan butir-butir soal yang bermutu yaitu soal-soal yang valid dan reliabel serta terbebas dari adanya tes yang mengandung bias butir yaitu butir-butir soal yang dapat menguntungkan atau merugikan pihak-pihak tertentu. Tes yang terbebas bias butir akan memberikan informasi yang benar terhadap karakteristik butir dan kemampuan peserta didik. Dengan demikian, tes terbebas dari bias butir dapat memberikan hasil yang dapat dipertanggungjawabkan (akuntabel).

b. Sekolah

Bagi sekolah-sekolah yang memiliki peserta didik dengan beragam latar belakang (etnis, kebudayaan, gender, sosial ekonomi, dan lain-lain) diharapkan memiliki pemahaman tentang tes yang mengandung bias butir. Hasil penelitian ini hendaknya dapat dimanfaatkan untuk mengembangkan tes yang terbebas bias butir, agar penilaian yang dilakukan merupakan penilaian yang objektif dan memenuhi prinsip-prinsip keadilan.

c. Dinas Pendidikan dan Pusat Penilaian Pendidikan Kemdikbud

Hasil penelitian ini dapat dijadikan informasi bagi dinas pendidikan dan Pusat Penilaian Pendidikan Kemdikbud untuk mengambil kebijakan terkait upaya-upaya peningkatan kemampuan guru untuk menyusun instrumen penilaian yang bermutu dan bebas bias butir. Sebagai *leading sector* dalam pengujian tingkat nasional (UN), Puspendik hendaknya memperhatikan berbagai keragaman yang ada di wilayah NKRI antara lain perbedaan suku, ras, agama, kebudayaan, adat-istiadat, dan perbedaan lainnya. Hasil penelitian ini diharapkan dapat memberikan informasi dan masukan untuk pengambilan kebijakan terkait dengan peningkatan mutu pengembangan bank soal yang akan dijadikan bahan untuk pelaksanaan UN.

BAB II

KAJIAN TEORETIK

Pada bagian ini akan disajikan dasar-dasar teoretis yang dijadikan acuan dalam penelitian yaitu konsep dasar bias butir dan metode *DIF* yang akan digunakan untuk mendeteksi bias butir. Metode *DIF* dalam penelitian ini menggunakan pendekatan teori responsi butir (*Item Response Theory/IRT*), oleh karena itu terlebih dahulu dipaparkan tentang Teori Responsi Butir, meliputi: (a) konsep dasar *IRT*, (b) asumsi-asumsi *IRT*, (c) parameter responden dan butir, (d) estimasi parameter, (e) model logistik, (f) pencocokan model, dan (g) fungsi informasi.

Sesuai dengan pendekatan yang digunakan dalam penelitian ini yaitu pendekatan teori responsi butir, maka metode pendeteksian bias butir juga dipersyaratkan memenuhi asumsi-asumsi dalam teori responsi butir. Pembahasan selanjutnya terkait dengan metode *DIF*, berturut-turut: (a) pengertian *DIF* dan bias butir, (b) jenis-jenis *DIF*, (c) bias gender, (d) metode *DIF* yang diperbandingkan yaitu metode Uji Beda Taraf Sukar, Khi-kuadrat Lord, dan Distribusi Sampling Empiris, serta (f) sensitivitas metode *DIF*.

Sebelum dilakukan proses analisis pendeteksian butir-butir yang mengandung *DIF*, terlebih dahulu dilakukan penyetaraan parameter butir dan responden untuk kelompok fokal dan referensi. Dengan demikian

pembahasan selanjutnya adalah penyetaraan parameter yang digunakan pada *linking*, dilanjutkan dengan kajian terhadap hasil-hasil penelitian yang relevan berupa hasil-hasil penelitian yang ada hubungannya dengan pembahasan metode *DIF*, dari berbagai jurnal dan disertasi yang melakukan penelitian sebelumnya. Hasil kajian secara teoretik dan empirik (hasil penelitian yang relevan) dijadikan dasar untuk merumuskan kerangka teoretik, dan pengajuan hipotesis penelitian.

A. Deskripsi Konseptual

Berikut ini dipaparkan dasar-dasar teoretik yang digunakan sebagai pedoman dalam melaksanakan penelitian. Teori-teori yang diangkat dalam penelitian ini berdasarkan Teori Responsi Butir.

1. Teori Responsi Butir (*Item Response Theory/IRT*)

Banyak istilah yang digunakan terkait dengan *Item Response Theory/IRT*. Teori Responsi Butir dikenal juga sebagai teori ciri laten (*Latent Trait Theory/LTT*) atau lengkungan karakteristik butir (*Item Characteristic Curve/ICC*) atau fungsi karakteristik butir (*Item Characteristic Function/ICF*).¹⁹ Walaupun sedemikian banyak istilah yang digunakan oleh para ahli untuk menyatakan pengertian (konsep) teori responsi butir, dalam penelitian ini,

¹⁹ Dali S. Naga, *op. cit*, hh. 160-161.

istilah yang akan digunakan adalah Teori Responsi Butir (*Item Response Theory/IRT*).

a. Konsep Dasar IRT

Secara garis besar, terdapat dua pendekatan dalam pengukuran, yaitu pendekatan teori tes klasik (*Classical Test Theory/CTT*) dan pendekatan teori tes modern (*Item Response Theory/IRT*).²⁰ Teori tes klasik telah berkembang secara luas pada bidang psikologi dalam kurun waktu yang cukup lama. Namun dalam perjalanan tersebut dijumpai beberapa kelemahan dalam teori tes klasik, antara lain nilai statistik butir yang sering digunakan dalam pengukuran seperti tingkat kesukaran (*difficulty*) dan daya pembeda (*discrimination*) tergantung pada karakteristik peserta tes. Bila kemampuan peserta tes tinggi maka butir soal menjadi mudah dan sebaliknya bila kemampuan peserta tes rendah maka butir soal memiliki tingkat kesukaran yang tinggi.²¹ Statistik lain yang juga tergantung pada kemampuan peserta tes adalah koefisien validitas dan reliabilitas. Model ketergantungan tersebut sering disebut sebagai *group dependent* dan *item dependent*.

Carlo Magno (2009) mengemukakan bahwa terdapat sejumlah keterbatasan pada teori tes klasik, dan banyak keunggulan pada *IRT*. Hasil

²⁰ Embretson, S. E., & Reise, S. P., *Item Response Theory for Psychologists* (NJ: Lawrence Erlbaum Associates Inc., 2000), h. 12.

²¹ Hambleton, R. K., dan Swaminathan H., *Item Response Theory* (Boston: Kluwer Nijhoff Publishing, 1985), hh. 1-2.

penelitiannya menyatakan: (1) nilai estimasi parameter taraf sukar butir pada pendekatan *IRT* tidak berubah pada sampel yang berbeda, sebaliknya pada pendekatan teori tes klasik nilai estimasi parameter taraf sukar berubah-ubah sesuai kemampuan peserta tes; (2) parameter taraf sukar juga lebih stabil pada pendekatan *IRT* dibandingkan dengan pendekatan teori tes klasik, pada bentuk tes yang berbeda-beda; (3) konsistensi internal *IRT* sangat stabil di antara sampel, sebaliknya konsistensi internal pendekatan teori tes klasik gagal mencapai kestabilan di antara sampel; dan (4) pendekatan *IRT* memiliki kesalahan pengukuran lebih kecil dibandingkan dengan pendekatan teori tes klasik.²²

Romel A. Morales (2009) menyatakan bahwa pendekatan teori tes klasik dan *IRT* merupakan pendekatan pengukuran yang memiliki *framework* yang sangat berbeda. Pendekatan teori tes klasik telah digunakan sangat lama, dan menjadi pendekatan pengukuran yang sangat familiar dan populer. Walaupun demikian, pendekatan teori tes klasik memiliki sejumlah kekurangan (keterbatasan) sehingga perlu penyempurnaan. Untuk itulah *IRT* ada, untuk mengatasi berbagai kelemahan dalam teori tes klasik.²³

²² Carlo Magno, Demonstrating the Difference between Classical Test Theory and Item Response Theory Using Derived Test Data. *The International Journal of Educational and Psychological Assessment* April 2009, Vol. 1, Issue 1, pp. 1-11 (diakses 5 Maret 2016), h. 1.

²³ Romel A. Morales, Evaluation of Mathematics Achievement Test: A Comparison between CTT and *IRT*. *The International Journal of Educational and Psychological Assessment* April 2009, Vol. 1, Issue 1, pp. 19-26 (diakses 5 Maret 2016), h. 19.

Teori tes modern *IRT* dikembangkan untuk mengatasi berbagai keterbatasan dalam teori tes klasik. Hasil yang diperoleh melalui pendekatan teori klasik dirasakan kurang akurat. Teori responsi butir berusaha meningkatkan akurasi pengukuran melalui pemisahan parameter butir dari kemampuan responden. Berapapun kemampuan responden, nilai parameter butir tidak berubah. Karakteristik butir ditentukan oleh responsi para responden (baik kemampuan tinggi maupun kemampuan rendah) sehingga dikenal sebagai teori responsi butir (*Item Response Theory*).²⁴

IRT sebagai solusi untuk mengatasi berbagai kelemahan dalam teori tes klasik memiliki beberapa karakteristik sebagai berikut: 1) parameter butir tidak tergantung pada parameter kemampuan peserta tes, 2) skor tes dapat menggambarkan kemampuan subjek, 3) model dinyatakan dalam bentuk tingkatan (level) butir, tidak dalam tingkatan tes, 4) model tidak memerlukan tes paralel untuk menghitung nilai koefisien reliabilitas, dan 5) model menyediakan ukuran yang tepat untuk setiap skor kemampuan.²⁵

Selanjutnya *IRT* didasarkan pada dua buah postulat yaitu: a) prestasi subjek pada salah satu butir soal dapat diprediksikan dengan seperangkat faktor yang disebut kemampuan laten (*latent traits*), dan b) hubungan antara

²⁴ Dali S. Naga, *Teori Sekor pada Pengukuran Mental* (Jakarta: PT. Nagarani Citrayasa, 2013), h. 457.

²⁵ Hambleton, R. K., Swaminathan, H., & Rogers, H. J., *Fundamentals of Item Response Theory* (CA: Sage Publication Inc., 1991), hh. 2-5.

prestasi subjek pada suatu butir dan kemampuan yang mendasarinya sesuai dengan fungsi grafik naik monoton tertentu, yang disebut kurva karakteristik butir (*Item Characteristic Curve/ICC*). Kurva karakteristik butir ini menggambarkan bahwa semakin tinggi level kemampuan peserta tes, semakin meningkat pula peluang menjawab benar suatu butir soal, dan sebaliknya semakin rendah level kemampuan peserta tes maka semakin rendah peluang menjawab benar suatu butir soal.²⁶

b. Asumsi-Asumsi dalam *IRT*

Model matematis dalam *IRT* mempunyai makna bahwa probabilitas peserta tes menjawab benar sebuah butir tertentu tergantung pada kemampuan peserta tes dan parameter butir. Hal ini berarti bahwa peserta tes dengan kemampuan tinggi, memiliki probabilitas menjawab dengan benar sebuah butir lebih besar daripada peserta tes dengan kemampuan rendah, dan sebaliknya. Hubungan antara kemampuan dan probabilitas menjawab benar digambarkan dalam bentuk kurva karakteristik butir.

Agar butir pada *IRT* mengikuti karakteristik butir, maka diperlukan sejumlah persyaratan. Syarat-syarat tersebut diperlukan untuk menjamin bahwa bentuk teori responsi butir sesuai dengan karakteristik butir yang digunakan. Terdapat 3 (tiga) asumsi dasar dalam *IRT* yaitu: unidimensi,

²⁶ *Ibid.*, hh. 2-5.

invariansi parameter, dan independensi lokal. Ketiga asumsi tersebut dapat dijelaskan sebagai berikut.²⁷

1) Unidimensi

Unidimensi artinya setiap butir tes hanya mengukur satu kemampuan. Teori responsi butir menghendaki lengkungan yang sesuai dengan karakteristik butir. Lengkungan itu terus naik mengikuti kenaikan pada parameter kemampuan. Agar lengkungan probabilitas jawaban betul terus naik menurut kenaikan parameter kemampuan, maka atribut yang diukur dibatasi hanya pada satu macam atau satu dimensi atribut. Pada satu dimensi atribut itu probabilitas jawaban betul terus meningkat pada peningkatan kemampuan responden.²⁸ Pada praktiknya, asumsi unidimensi tidak dapat dipenuhi secara ketat karena adanya faktor-faktor kognitif, kepribadian, kecemasan, motivasi, dan tendensi untuk menebak. Oleh karena itu, asumsi unidimensi dapat ditunjukkan hanya jika tes mengandung satu komponen dominan yang mengukur prestasi responden.²⁹

Lebih lanjut Dali S. Naga (1992) mengemukakan bahwa persyaratan unidimensi ditujukan untuk mempertahankan invariansi pada teori responsi butir. Apabila sebuah butir sampai mengukur lebih dari satu dimensi, maka

²⁷ Hambleton, R. K., Swaminathan H., dan Roger H. J., *op. cit.*, hh. 9-19.

²⁸ Dali S. Naga, *op. cit.*, hh. 459-460.

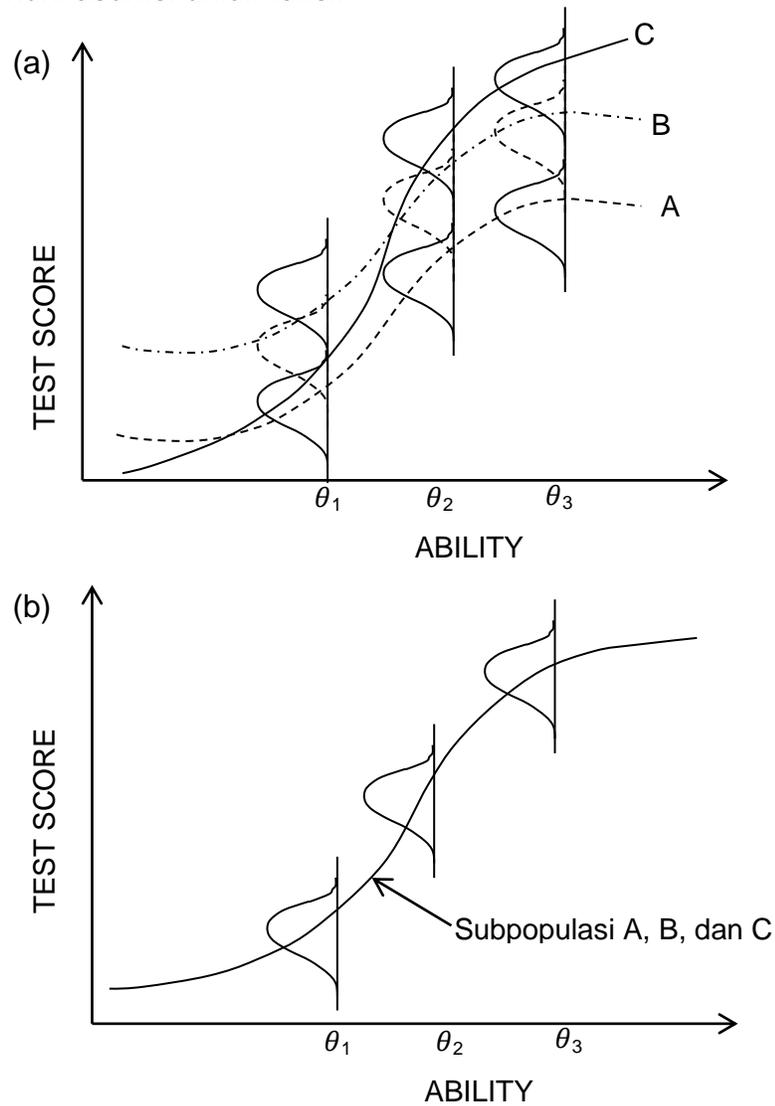
²⁹ Heri Retnawati, *op. cit.*, h. 1.

jawaban terhadap butir tersebut akan merupakan kombinasi dari berbagai kemampuan pada peserta. Akibatnya, tidak dapat diketahui kontribusi dari setiap kemampuan terhadap jawaban peserta tersebut. Dengan demikian, apabila mengganti butir tes atau mengganti kelompok peserta mengakibatkan invariansi pada ukuran ciri butir dan ciri peserta tidak dapat dipertahankan.³⁰

Hubungan antara kemampuan responden dan skor tes yang dicapai dinyatakan dalam kurva yang tidak linear. Asumsi bahwa pengukuran *latent trait* harus memenuhi salah satu syarat unidimensional, dilakukan ketika akan mengkonstruksi tes yang bertujuan untuk meningkatkan *interpretability* dari satu set skor tes. Memperhatikan distribusi kondisional dari nilai tes pada tingkat kemampuan tertentu untuk beberapa subpopulasi, gambar 2.1 menyajikan ilustrasi suatu distribusi kondisional pada suatu bagian level kemampuan subpopulasi peserta tes. Di sepanjang garis regresi, terdapat sebaran skor tes. Variabilitas kesalahan pengukuran skor tes mungkin terjadi. Misalkan ada 3 subpopulasi A, B, dan C. Jika distribusi bervariasi pada masing-masing subpopulasi A, B, dan C, maka tes tidak hanya mengukur kemampuan tunggal saja. Pada gambar 2.1a terlihat bahwa masing-masing subpopulasi A, B, dan C memiliki distribusi yang berbeda. Sebaliknya, jika skor tes disajikan sebagai grafik fungsi yang sama, maka tes mengukur satu dimensi kemampuan saja. Pada gambar 2.1b, terlihat bahwa ketiga

³⁰ Dali S. Naga, op. cit., hh. 164-165

subpopulasi memiliki satu distribusi saja, menunjukkan bahwa butir soal memenuhi asumsi unidimensi.³¹



Gambar 2.1 Distribusi Kondisional Skor Tes 3 Level Kemampuan, pada 3 Subpopulasi A, B, C.

³¹ Hambleton, R. K., dan Swaminathan H., *op. cit.*, hh. 18-19.

2) Invariansi Parameter

Invariansi parameter artinya karakteristik butir soal tidak tergantung pada distribusi parameter kemampuan peserta tes, dan parameter yang menjadi ciri peserta tes tidak tergantung pada ciri butir soal. Kemampuan seseorang tidak akan berubah hanya karena mengerjakan tes yang berbeda tingkat kesukarannya, dan parameter butir soal tidak akan berubah bila diujikan pada kelompok responden yang berbeda kemampuannya.³²

Invariansi parameter kemampuan dapat diselidiki dengan mengajukan dua perangkat tes atau lebih yang memiliki tingkat kesukaran yang berbeda pada sekelompok peserta tes. Invariansi parameter kemampuan akan terbukti jika hasil estimasi kemampuan peserta tes tidak berbeda walaupun tes yang dikerjakan berbeda tingkat kesukarannya. Invariansi parameter butir terbukti jika hasil estimasi parameter butir tidak berbeda walaupun diujikan pada kelompok peserta yang berbeda kemampuannya.

3) Independensi Lokal

Hambleton, R. K., Swaminathan, H., & Rogers (1991) mengemukakan bahwa independensi lokal artinya parameter kemampuan peserta tes yang satu tidak tergantung pada kemampuan peserta tes yang lainnya, demikian pula parameter butir yang satu tidak tergantung pada parameter butir soal

³² Heri Retnawati, "*Estimasi Efisiensi Relatif Tes Berdasarkan Teori Responsi Butir dan teori Tes Klasik*" (Disertasi: Universitas Negeri Yogyakarta, 2008), h. 34.

yang lainnya. Independensi lokal adalah independensi secara statistika. Untuk menguji independensi lokal dapat dilakukan dengan membuktikan bahwa peluang dari pola jawaban setiap peserta tes sama dengan hasil kali peluang jawaban peserta tes pada setiap butir soal. Secara matematis, independensi lokal dapat dinyatakan dalam bentuk sebagai berikut.³³

$$P(u_1, u_2, u_3, \dots, u_n|\theta) = P(u_1|\theta) \cdot P(u_2|\theta) \cdot P(u_3|\theta) \dots P(u_n|\theta) \\ = \prod_{i=1}^n P(u_i|\theta) \dots\dots\dots (2.1)$$

Keterangan:

- i : butir soal ke- i , $i=1, 2, 3, \dots, n$
 n : banyak butir soal
 $P(u_i|\theta)$: probabilitas peserta tes dengan kemampuan θ , menjawab benar butir soal ke- i
 $P(u_1, u_2, u_3, \dots, u_n|\theta)$: probabilitas peserta tes dengan kemampuan θ , menjawab benar butir soal ke- i sampai ke- n .

c. Parameter Responden dan Butir

Parameter Responden. Dalam menjawab atau menanggapi suatu butir, responden memiliki kemampuan. Ada responden yang memiliki kemampuan tinggi sehingga memperoleh skor tinggi, serta ada pula responden yang memiliki kemampuan rendah sehingga memperoleh skor rendah. Kemampuan responden ini dinyatakan sebagai parameter

³³ Hambleton, R. K., Swaminathan, H., & Rogers, H. J., *op. cit.*, h. 10.

kemampuan yang diberi lambang θ . Parameter kemampuan merupakan karakteristik responden. Pada model karakteristik butir, proporsi jawaban betul dinyatakan sebagai probabilitas jawaban betul dilambangkan dengan $P(\theta)$.³⁴ Nilai $P(\theta)$ dapat dihitung menggunakan rumus matematis yang disesuaikan dengan model yang digunakan.

Parameter Butir. Pada teori responsi butir, parameter butir terdiri dari: taraf sukar butir, daya pembeda, dan kebetulan menjawab betul. Masing-masing parameter tersebut dapat dijelaskan sebagai berikut.

- Taraf sukar butir (*difficulty*), dilambangkan dengan huruf b . Probabilitas jawaban betul ditentukan oleh selisih antara kemampuan responden dan taraf sukar butir ($\theta - b$). Semakin sukar suatu butir makin rendah probabilitas jawaban betul $P(\theta)$ dan sebaliknya.
- Daya pembeda (*discrimination*), dilambangkan dengan huruf a . Daya pembeda berkenaan dengan kepekaan suatu butir terhadap perbedaan di antara kemampuan responden. Semakin besar nilai a , berarti semakin peka butir tersebut terhadap perbedaan kemampuan responden. Dikaitkan dengan parameter taraf sukar butir dan parameter kemampuan ($\theta - b$), kepekaan parameter daya beda butir merupakan koefisien arah sehingga kedudukan parameter daya beda butir di dalam karakteristik butir terletak pada bentuk $a(\theta - b)$.

³⁴ Dali S. Naga, *op. cit.*, hh. 417-418.

- Kebetulan menjawab betul (*pseudoguessing*), dilambangkan dengan huruf c. Pada butir soal bentuk pilihan ganda, seorang responden mungkin saja menjawab benar butir tersebut karena terkaan semata. Probabilitas jawaban betul karena terkaan ini disebut sebagai parameter kebetulan menjawab betul.

d. Estimasi Parameter

Estimasi parameter merupakan cara untuk menentukan nilai parameter kemampuan responden dan parameter butir berdasarkan data respon peserta terhadap sejumlah butir soal. Secara umum estimasi parameter dalam teori responsi butir dibedakan menjadi dua yaitu estimasi secara terpisah (*sparate estimation*) dan estimasi secara serentak (*concurrent estimation*).

Proses estimasi terpisah dilakukan apabila salah satu di antara parameter kemampuan responden atau parameter butir telah diketahui. Apabila parameter butir telah diketahui maka estimasi terpisah akan mengestimasi parameter kemampuan responden. Sedangkan pada estimasi secara serentak, semua parameter diestimasi sekaligus. Estimasi serentak menghasilkan parameter kemampuan dan parameter butir.

Berikut ini disajikan langkah-langkah melakukan estimasi parameter kemampuan dan parameter butir model L3P secara terpisah, menggunakan metode *maximum likelihood*.

1) Jawaban peserta ke-g terhadap butir ke-i dinyatakan sebagai X_{gi} , dengan:

$$X_{gi} = \begin{cases} 1 & \text{bila jawaban benar} \\ 0 & \text{bila jawaban salah} \end{cases}$$

2) Probabilitas jawaban benar peserta ke-g terhadap butir ke-i dinyatakan sebagai $P_{gi}(\theta)$ serta probabilitas jawaban salah dinyatakan sebagai $Q_{gi}(\theta)$, dengan $Q_{gi}(\theta) = 1 - P_{gi}(\theta)$ (2.2)

3) Fungsi kebolehjadian dari jawaban peserta ke-g terhadap butir ke-i dinyatakan sebagai:

$$L(X_{gi}) = \prod_{g=1}^M \prod_{i=1}^N P_{gi}(\theta)^{X_{gi}} \cdot Q_{gi}(\theta)^{1-X_{gi}} \dots\dots\dots (2.3)$$

4) Untuk menyederhanakan perhitungan, maka bentuk perkalian tersebut diubah menjadi penjumlahan menggunakan logaritma, sehingga menjadi:

$$\ln L(X_{gi}) = \sum_{g=1}^M \sum_{i=1}^N (X_{gi} \ln P_{gi}(\theta) + (1 - X_{gi}) \ln Q_{gi}(\theta)). \dots\dots\dots (2.4)$$

5) Kebolehjadian maksimum akan dicapai pada nilai masing-masing parameter yang memenuhi persamaan berikut.

a) Parameter kemampuan (θ):

- $\frac{\partial \ln L}{\partial \theta} = 0$

$$D \sum_{i=1}^N a_i \frac{(X_i - P_i(\theta))(P_i(\theta) - c_i)}{P_i(\theta)(1 - c_i)} = 0 \dots\dots\dots (2.5)$$

- $\frac{\partial^2 \ln L}{\partial \theta^2} = D^2 \sum_{i=1}^N \frac{a_i^2 [P_i(\theta) - c_i] [X_i c_i - P_i^2(\theta)] Q_i(\theta)}{P_i^2(\theta) [1 - c_i]^2} \dots\dots\dots (2.6)$

Selanjutnya persamaan di atas diselesaikan dengan iterasi metode Newton-Raphson:

- $$\theta_{s+1} = \theta_s - \frac{\frac{\partial \ln L}{\partial \theta}}{\frac{\partial^2 \ln L}{\partial \theta^2}}$$
- $$\theta_{s+1} = \theta_s - \frac{\sum_{i=1}^N \frac{a_i [X_i - P_i(\theta)] [P_i(\theta) - c_i]}{P_i(\theta)(1-c_i)}}{D \sum_{i=1}^N \frac{a_i^2 [P_i(\theta) - c_i] [X_i c_i - P_i^2(\theta)] Q_i(\theta)}{P_i^2(\theta) [1-c_i]^2}} \dots\dots\dots (2.7)$$

Nilai awal (θ_0) yang digunakan adalah logit sukses: $\theta_0 = \ln \frac{P_i(\theta)}{Q_i(\theta)}$

b) Parameter tingkat kesukaran (b):

- $$\frac{\partial \ln L}{\partial b_i} = 0$$
- $$\sum_{g=1}^M - \frac{Da}{(1-c)} \left[\frac{[X_g - P_g(\theta)] [P_g(\theta) - c]}{P_g(\theta)} \right] = 0 \dots\dots\dots (2.8)$$

- $$\frac{\partial^2 \ln L}{\partial \theta^2} = \sum_{g=1}^M - \frac{D^2 a^2}{(1-c)^2} \frac{Q_g(\theta) [P_g(\theta) - c] [X_g c - P_g^2(\theta)]}{P_g^2(\theta)} \dots\dots\dots (2.9)$$

Selanjutnya diselesaikan dengan iterasi metode Newton-Raphson:

- $$b_{s+1} = b_s - \frac{\frac{\partial \ln L}{\partial b}}{\frac{\partial^2 \ln L}{\partial b^2}}$$

$$\bullet \quad b_{s+1} = b_s - \frac{\sum_{g=1}^M \frac{Da}{(1-c)} \left[\frac{[X_g - P_g(\theta)][P_g(\theta) - c]}{P_g(\theta)} \right]}{\sum_{g=1}^M \frac{D^2 a^2}{(1-c)^2} \frac{Q_g(\theta)[P_g(\theta) - c][X_g c - P_g^2(\theta)]}{P_g^2(\theta)}} \dots\dots\dots (2.10)$$

Nilai awal (b_0) yang digunakan adalah logit gagal: $b_0 = \ln \frac{Q_g(\theta)}{P_g(\theta)}$

c) Parameter daya pembeda (a):

$$\bullet \quad \frac{\partial \ln L}{\partial a_i} = 0$$

$$\sum_{g=1}^M \frac{D}{(1-c)} \left[\frac{[\theta_g - b][X_g - P_g(\theta)][P_g(\theta) - c]}{P_g(\theta)} \right] = 0 \dots\dots\dots (2.11)$$

$$\bullet \quad \frac{\partial^2 \ln L}{\partial a^2} = \sum_{g=1}^M - \frac{D^2}{(1-c)^2} \frac{[\theta_g - b]^2 Q_g(\theta)[P_g(\theta) - c][X_g c - P_g^2(\theta)]}{P_g^2(\theta)} \dots\dots\dots (2.12)$$

Selanjutnya diselesaikan dengan iterasi metode Newton-Raphson:

$$\bullet \quad a_{s+1} = a_s - \frac{\frac{\partial \ln L}{\partial a}}{\frac{\partial^2 \ln L}{\partial a^2}}$$

$$\bullet \quad a_{s+1} = a_s - \frac{\sum_{g=1}^M \frac{D}{(1-c)} \left[\frac{[\theta_g - b][X_g - P_g(\theta)][P_g(\theta) - c]}{P_g(\theta)} \right]}{\sum_{g=1}^M \frac{D^2}{(1-c)^2} \frac{[\theta_g - b]^2 Q_g(\theta)[P_g(\theta) - c][X_g c - P_g^2(\theta)]}{P_g^2(\theta)}} \dots\dots\dots (2.13)$$

d) Parameter *guessing* (c):

$$\bullet \quad \frac{\partial \ln L}{\partial c_i} = 0$$

$$\sum_{g=1}^M \frac{1}{(1-c)} \left[\frac{[X_g - P_g(\theta)]}{P_g(\theta)} \right] = 0 \dots\dots\dots (2.14)$$

$$\bullet \frac{\partial^2 \ln L}{\partial c^2} = \sum_{g=1}^M \frac{1}{(1-c)^2} \left[\left(\frac{X_g}{P_g(\theta)} - 1 \right) - \frac{X_g Q_g(\theta)}{P_g^2(\theta)} \right] \dots\dots\dots (2.15)$$

Selanjutnya diselesaikan dengan iterasi metode Newton-Raphson:

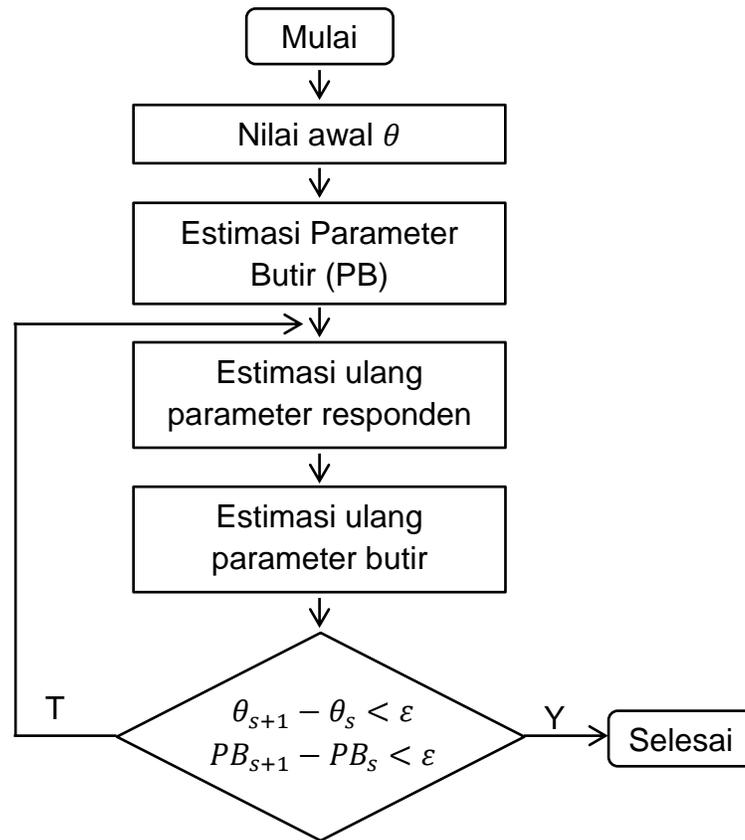
$$\bullet c_{s+1} = c_s - \frac{\frac{\partial \ln L}{\partial c}}{\frac{\partial^2 \ln L}{\partial c^2}}$$

$$\bullet c_{s+1} = c_s - \frac{\sum_{g=1}^M \frac{1}{(1-c)} \left[\frac{[X_g - P_g(\theta)]}{P_g(\theta)} \right]}{\sum_{g=1}^M \frac{1}{(1-c)^2} \left[\left(\frac{X_g}{P_g(\theta)} - 1 \right) - \frac{X_g Q_g(\theta)}{P_g^2(\theta)} \right]} \dots\dots\dots (2.16)$$

6) Proses iterasi dihentikan apabila selisih nilai estimasi parameter $|\xi_{s+1} - \xi_s| < \varepsilon$, dengan ε bilangan yang sangat kecil dan ξ adalah parameter kemampuan dan butir. Nilai ξ_{s+1} merupakan nilai maksimum *likelihood* estimasi parameter parameter kemampuan dan butir.

Untuk melakukan estimasi secara serentak menggunakan metode *maximum likelihood* melalui iterasi pada model L3P, memerlukan banyak langkah. Bahkan langkah iterasi itu berlangsung bolak balik di antara parameter. Langkah-langkah iterasi secara serentak menggunakan metode *maximum likelihood* dapat digambarkan sebagai berikut.³⁵

³⁵ Dali S. Naga, *op. cit.*, h. 518.



Gambar 2.2 Peta Alir Langkah Iterasi

Berdasarkan gambar 2.2, dapat dijelaskan langkah-langkah iterasi secara serentak menggunakan metode *maximum likelihood* adalah sebagai berikut.

- 1) Tentukan nilai awal θ_0 , yang diambil dari nilai logit sukses. Dengan nilai awal θ_0 ini, melalui iterasi ditemukan parameter butir PB_0 .
- 2) Dengan nilai parameter butir PB_0 , melalui iterasi, ditemukan parameter θ_1 . Bila selisih θ_1 dan θ_0 masih besar, maka iterasi terus dilanjutkan.
- 3) Dengan nilai θ_1 , melalui iterasi diperoleh parameter butir PB_1 . Jika selisih PB_1 dan PB_0 cukup besar, maka parameter butir terus diiterasi.

- 4) Demikian seterusnya iterasi berlangsung bolak balik di antara parameter θ dan parameter PB. Iterasi dihentikan jika selisih di antara parameter berurutan sudah cukup kecil $\theta_{s+1} - \theta_s < \varepsilon$ dan $PB_{s+1} - PB_s < \varepsilon$. Misalnya $\varepsilon \approx 0,01$.
- 5) Rumus-rumus untuk menentukan θ_{s+1} dan PB_{s+1} yang digunakan dalam perhitungan iterasi Newton-Raphson, dapat dilihat pada rumus-rumus (2.7), (2.10), (2.13), dan (2.16).

e. Model Logistik

Dali S. Naga (2013) mengemukakan bahwa model logistik diambil dari distribusi probabilitas logistik. Bentuk umum fungsi densitas distribusi probabilitas logistik adalah sebagai berikut.³⁶

$$f(x) = \frac{e^{\frac{x-u}{k}}}{k\left(1+e^{\frac{x-u}{k}}\right)^2} = \frac{\operatorname{sech}^2 \frac{x-u}{2k}}{4k} \dots\dots\dots (2.17)$$

Seperti halnya pada karakteristik butir normal, bentuk yang digunakan adalah ojaif atau fungsi distribusi. Fungsi distribusi pada bentuk distribusi probabilitas logistik adalah sebagai berikut.

$$\Psi(x) = \int_{-\infty}^x \frac{e^{\frac{x-u}{k}}}{k\left(1+e^{\frac{x-u}{k}}\right)^2} dx = \frac{e^{\frac{x-u}{k}}}{1+e^{\frac{x-u}{k}}} \dots\dots\dots (2.18)$$

³⁶ Dali S. Naga, *op. cit.*, hh. 436-437.

Pada karakteristik butir normal, digunakan distribusi probabilitas normal baku dengan rerata sama dengan 0, dan simpangan baku sama dengan 1. Oleh karena itu, bentuk khusus distribusi probabilitas logistik adalah $u=0$ dan $k=1$, sehingga bentuk distribusi probabilitas logistik sebagai berikut.

$$\Psi(x) = \frac{e^x}{1+e^x} \dots\dots\dots (2.19)$$

Hubungan antara kemampuan peserta tes terhadap harapan menjawab benar, dinyatakan dalam bentuk probabilitas. Dalam *IRT* model distribusi yang digunakan adalah distribusi logistik bukan distribusi normal. Hal ini disebabkan karena kurva normal berbentuk lonceng, sehingga kurva tidak monoton naik.³⁷ Keadaan tersebut mengakibatkan bahwa suatu kemampuan yang lebih tinggi daripada rerata, nilai probabilitasnya justru lebih rendah bila dibandingkan dengan probabilitas rerata kemampuan. Hal ini bertentangan dengan prinsip pengukuran bahwa peserta dengan kemampuan tinggi mempunyai peluang yang tinggi pula untuk menjawab benar butir soal tertentu. Pada penghitungan luas daerah di bawah kurva, dapat dilakukan dengan integral karena merupakan fungsi kerapatan peluang kontinu. Dengan adanya variabel kemampuan yang dikuadratkan pada fungsi

³⁷ Walpole, R. E., Mers, R. H., Myers, S. R., *Probability and Statistic for Engineers and Scientist* (Upper Saddle River, NJ: Prentice Hall, 2002), h. 31.

kerapatan peluang normal menyebabkan integral menjadi lebih rumit. Hal ini yang menyebabkan digunakannya model logistik pada teori responsi butir.³⁸

Terdapat 3 (tiga) model logistik dalam teori responsi butir, yaitu: a) model logistik 1 (satu) parameter (L1P), b) model logistik 2 (dua) parameter (L2P), dan c) model logistik 3 (tiga) parameter (L3P). Perbedaan di antara ketiga model tersebut terletak pada banyaknya parameter yang digunakan untuk menggambarkan karakteristik butir pada model yang digunakan. Parameter-parameter yang digunakan adalah taraf sukar butir b , daya pembeda a , dan indeks tebakan semu c .

1) Model Logistik L1P

Pada model logistik satu parameter, probabilitas peserta tes menjawab benar butir soal tertentu ditentukan oleh satu parameter butir yaitu tingkat kesukaran butir b_i . Secara teoretik nilai b_i terletak di antara $-\infty$ dan $+\infty$, tetapi secara praktis berkisar antara -2 dan +2. Butir soal dengan $b_i=-2$ merupakan butir soal yang sangat mudah, sedangkan butir soal dengan nilai $b_i=+2$ merupakan butir soal yang sangat sukar. Secara matematis, model logistik satu parameter dapat dinyatakan sebagai berikut.³⁹

$$P_i(\theta) = \frac{e^{D(\theta-b_i)}}{1+e^{D(\theta-b_i)}} \text{ atau } P_i(\theta) = \frac{1}{1+e^{-D(\theta-b_i)}} \dots\dots\dots (2.20)$$

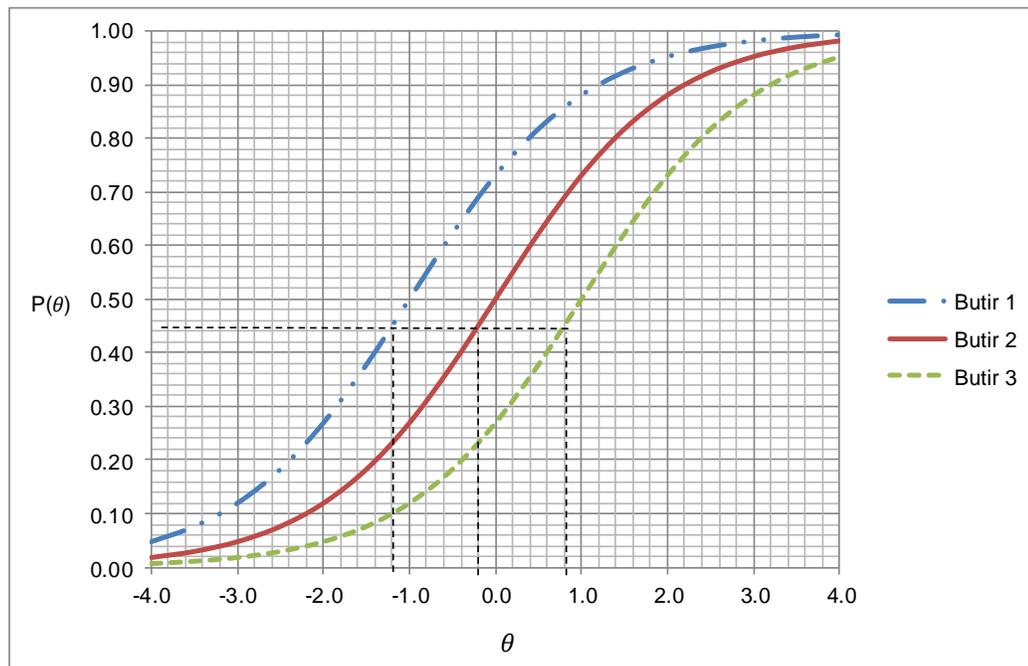
³⁸ Hog, R. V., & Craig, A. T., *Introduction to Mathematical Statistics 4th ed* (New York: Coller Macmillan, 1978), h. 92.

³⁹ Hambleton, R. K., Swaminathan, H., & Rogers, H. J., *op. cit.*, h. 12.

Keterangan:

- i : butir soal ke- i , $i=1, 2, 3, \dots, n$
- e : bilangan natural besarnya mendekati 2,718
- b_i : tingkat kesukaran butir soal ke- i
- θ : parameter kemampuan
- D : konstanta besarnya 1,7
- $P_i(\theta)$: probabilitas peserta tes dengan kemampuan θ , menjawab benar butir soal ke- i

Parameter b_i merupakan suatu titik pada skala kemampuan (θ) dalam kurva karakteristik butir (*item characteristic curve/ICC*) ketika peluang menjawab benar peserta tes sebesar 50%. Misalkan suatu butir tes memiliki tingkat kesukaran $b_i = 0,4$ berarti diperlukan kemampuan (θ) minimal 0,4 pada skala kemampuan untuk dapat menjawab dengan benar suatu butir dengan peluang 50%. Semakin besar nilai parameter b_i , semakin besar pula kemampuan yang dibutuhkan untuk bisa menjawab dengan benar suatu butir soal tertentu dengan peluang 50%. Artinya semakin besar nilai b_i semakin sulit butir soal tersebut. Berikut ini disajikan ilustrasi kurva karakteristik butir, yang terdiri atas tiga butir soal dengan tingkat kesukaran untuk butir soal nomor 1, 2, dan 3 berturut-turut $b_1= -1,0$; $b_2=0,0$; dan $b_3=1,0$.



Gambar 2.3 Kurva Karakteristik Butir Model L1P

2) Model Logistik L2P

Pada model logistik dua parameter, probabilitas peserta tes menjawab benar butir soal tertentu ditentukan oleh dua parameter butir yaitu tingkat kesukaran butir b_i dan daya pembeda a_i . Pada kurva karakteristik butir, a_i merupakan koefisien arah garis singgung (*slope*) pada titik $\theta = b_i$. Butir soal yang memiliki nilai a_i besar maka kurva karakteristik butir akan curam, sebaliknya butir soal yang memiliki nilai a_i kecil maka kurva karakteristik butir akan landai. Pada butir soal yang baik, nilai a_i mempunyai hubungan yang positif antara parameter butir dengan kemampuan peserta tes. Butir soal yang memiliki nilai a_i tinggi, menyatakan butir soal tersebut peka terhadap

perbedaan kemampuan peserta tes. Kepekaan ini tampak pada nilai probabilitas jawaban betul. Pada butir yang peka, sedikit saja kemampuan responden berbeda mengakibatkan cukup besar perbedaan nilai probabilitas jawaban betul. Nilai a_i berkisar antara 0 sampai 2. Secara matematis, model logistik dua parameter dapat dinyatakan sebagai berikut.⁴⁰

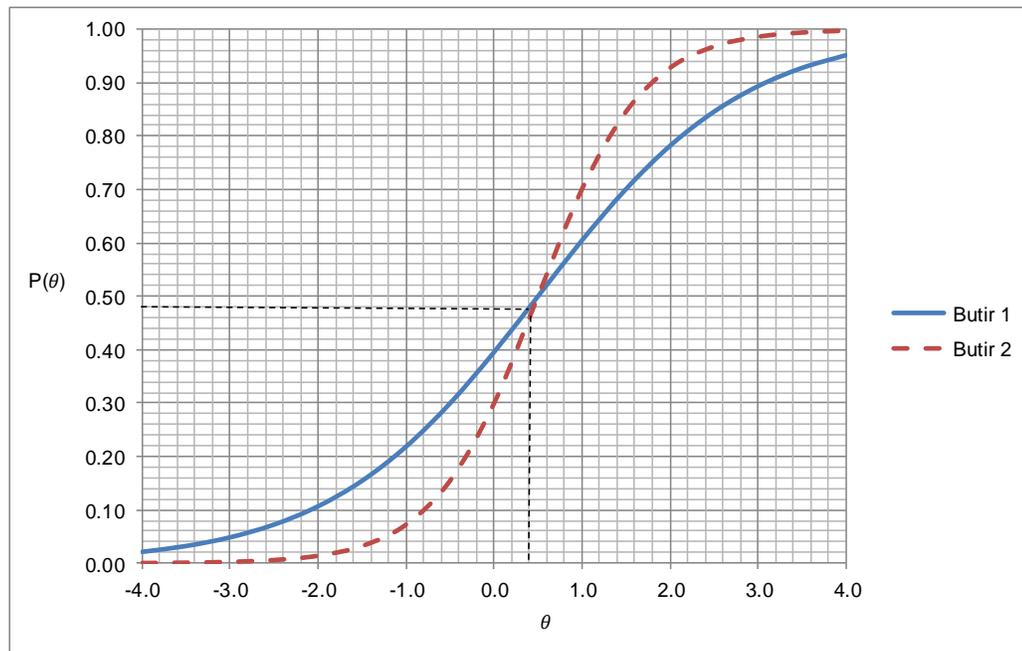
$$P_i(\theta) = \frac{e^{Da_i(\theta-b_i)}}{1+e^{Da_i(\theta-b_i)}} \text{ atau } P_i(\theta) = \frac{1}{1+e^{-Da_i(\theta-b_i)}} \dots\dots\dots (2.21)$$

Keterangan:

- i : butir soal ke-i, $i=1, 2, 3, \dots, n$
- e : bilangan natural besarnya mendekati 2,718
- b_i : tingkat kesukaran butir soal ke-i
- a_i : daya pembeda butir soal ke-i
- D : konstanta besarnya 1,7
- $P_i(\theta)$: probabilitas peserta tes dengan kemampuan θ , menjawab benar butir soal ke-i

Berikut ini disajikan ilustrasi kurva karakteristik butir model L2P, pada dua butir soal dengan tingkat kesukaran dan daya pembeda untuk butir soal nomor 1 dengan $a_1=0,5$ dan $b_1=0,5$; dan butir nomor 2 dengan $a_2=1$ dan $b_2=0,5$.

⁴⁰ Hambleton, R. K., Swaminathan, H., *op. cit.*, hh. 36-37.



Gambar 2.4 Kurva Karakteristik Butir Model L2P

Pada gambar 2.4 tampak grafik butir soal nomor 2 lebih curam daripada grafik butir soal nomor 1. Hal itu berarti bahwa soal nomor 2 lebih peka daripada soal nomor 1, karena nilai $a_2 > a_1$.

3) Model Logistik L3P

Pada model logistik tiga parameter, probabilitas peserta tes menjawab benar butir soal tertentu ditentukan oleh tiga parameter butir yaitu tingkat kesukaran butir b_i , daya pembeda a_i , dan indeks tebakan semu (*pseudo guessing*) c_i . Dengan adanya tebakan semu pada model logistik tiga parameter, memungkinkan peserta tes dengan kemampuan rendah mempunyai peluang lebih besar untuk menjawab dengan benar butir soal

tertentu. Parameter c_i akan menjadi asimtot bawah yang tidak nol (*non zero lower asymtot*) pada kurva karakteristik butir. Parameter ini menggambarkan probabilitas peserta tes dengan kemampuan rendah menjawab benar butir soal tertentu yang mempunyai tingkat kesukaran yang tidak sesuai dengan kemampuan peserta tes tersebut. Besarnya nilai c_i berkisar antara 0 dan 1. Sebuah butir soal dikatakan baik bila nilai c_i tidak lebih dari $1/k$, dengan k adalah banyaknya alternatif jawaban yang disediakan.⁴¹ Bila sebuah butir soal pilihan ganda terdiri atas 5 pilihan jawaban, maka butir soal tersebut dikatakan baik bilai nilai c_i tidak lebih dari 0,2. Secara matematis, model logistik tiga parameter dapat dinyatakan sebagai berikut.⁴²

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{Da_i(\theta-b_i)}}{1+e^{Da_i(\theta-b_i)}} \dots\dots\dots (2.22)$$

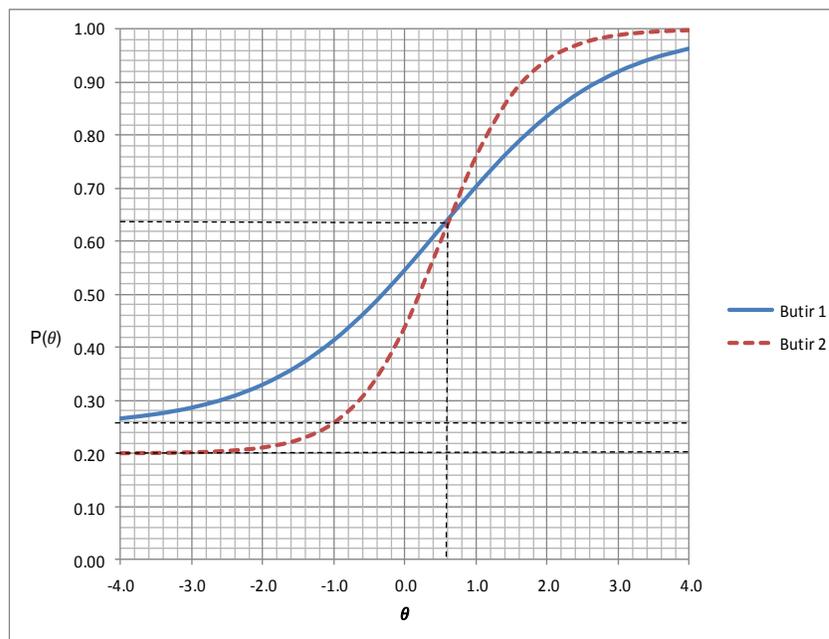
Keterangan:

- i : butir soal ke- i , $i=1, 2, 3, \dots, n$
- e : bilangan natural besarnya mendekati 2,718
- b_i : tingkat kesukaran butir soal ke- i
- a_i : daya pembeda butir soal ke- i
- c_i : indeks tebakan semu
- D : konstanta besarnya 1,7
- $P_i(\theta)$: probabilitas peserta tes dengan kemampuan θ , menjawab benar butir soal ke- i

⁴¹ Hulin, C.L., Drasgow, F., & Parsons, C. K., *Item Response Theory: Application to Psychological Measurement* (Homewood, IL: Dow Jornes-Irwin, 1983), h. 36.

⁴² Hambleton, R. K., Swaminathan, H., *op. cit.*, hh. 37-38.

Berikut ini disajikan ilustrasi kurva karakteristik butir model L3P, pada dua butir soal dengan tingkat kesukaran, daya pembeda, dan indeks tebakan semu sebagai berikut: butir soal nomor 1 dengan $a_1=0,5$; $b_1=0,5$; $c_1=0,25$; dan butir nomor 2 dengan $a_2=1$; $b_2=0,5$; $c_2=0,2$. Pada gambar 2.4 terlihat bahwa kurva butir soal nomor 2 lebih curam bila dibandingkan dengan kurva butir soal nomor 1, karena $a_2 > a_1$. Asimtot bawah kurva butir soal nomor 1 adalah 0,25, dengan demikian kemampuan peserta tes yang semestinya 0,5 bergeser ke atas sejauh 0,25 sehingga menjadi 0,75. Demikian pula untuk butir soal nomor 2, asimtot bawah kurva menjadi 0,2 sehingga kemampuan peserta tes yang semestinya 0,5 bergeser ke atas sejauh 0,2 sehingga menjadi 0,7.



Gambar 4.5 Kurva Karakteristik Butir Model L3P

f. Pencocokan Model

Dalam *IRT*, selain asumsi-asumsi yang telah diuraikan sebelumnya, hal penting yang perlu diperhatikan adalah pencocokan model. Pemilihan model yang tepat akan mengungkapkan keadaan yang sesungguhnya dari data tes sebagai hasil pengukuran.

Pencocokan model dilakukan dengan tujuan untuk memeriksa kecocokan antara model karakteristik yang dipilih dengan data dari lapangan. Dalam pemeriksaan ini kemungkinan ada butir yang cocok dan ada butir yang tidak cocok. Apabila model tersebut ditetapkan maka semua perhitungan akan berkisar di sekitar model itu. Sehingga bila data lapangan cocok dengan model karakteristik butir yang dipilih, maka estimasi terhadap parameter dapat dilakukan dengan akurat.

Data lapangan yang cocok dengan model karakteristik butir yang digunakan adalah data yang memenuhi ketentuan model. Pada teori responsi butir, setiap butir adalah independen sehingga kecocokan model dilakukan terhadap butir demi butir. Hasil pemeriksaan lain yang diperoleh adalah kecocokan estimasi parameter kemampuan dengan data dari lapangan. Dalam teori responsi butir karakteristik butir dan peserta adalah invarian. Sehingga hasil pengukuran yang tepat akan dapat mengungkap karakteristik butir dan peserta dengan objektif.

Dali S. Naga (1992) mengemukakan bahwa untuk mencocokkan model terdapat beberapa cara. Adapun cara-cara yang sering digunakan oleh para ahli, akan diuraikan secara rinci sebagai berikut.⁴³

1) Pencocokan secara statistika

Secara statistika, kecocokan antara data lapangan yang diperoleh dengan model dilihat dari lengkungan responsi butir yang dihasilkan oleh model yang dipilih. Menurut Naga (1992), cara statistika ini mengandung banyak kelemahan, karena cara ini terlalu peka terhadap ukuran sampel. Bila ukuran sampel peserta besar, cara ini cenderung menolak butir yang menyatakan butir tidak cocok dengan model sekalipun butir itu masih cukup memadai. Sebaliknya pada ukuran sampel peserta yang kecil, cara ini cenderung menerima butir yakni menyatakan butir itu cocok dengan model, sekalipun butir itu tidak memadai.

2) Pencocokan secara pemenuhan persyaratan pada model

Pada cara ini, pengujian kecocokan model dilihat dari apakah semua persyaratan untuk model itu telah dipenuhi oleh data yang terkumpul. Pada model logistik, persyaratan itu mencakup: unidimensi, bukan tes kecepatan; pada model L2P tidak terdapat jawaban terkaan, atau walaupun ada, jawaban terkaan adalah sekecil-kecilnya; dan pada model L1P nilai parameter daya pembeda sama untuk setiap butir.

⁴³ Dali S. Naga, *op. cit.*, hh. 288-289.

3) Pencocokan melalui pemenuhan harapan yang diinginkan

Harapan yang ingin dicapai pada model ini berisikan invariansi sehingga pencocokan akan memeriksa seberapa jauh invariansi yang ditentukan oleh teori responsi butir itu tercapai.

4) Pencocokan melalui kecermatan pada prediksi model

Cara pencocokan model ini dilakukan dengan melihat hasil prediksi berupa data sesungguhnya serta hasil prediksi pada data simulasi dan seberapa besar selisihnya. Pada pencocokan ini diharapkan adanya kecermatan tinggi atau selisih di antara kedua prediksi di atas sangat kecil.

Dilihat dari uraian di atas, terdapat beberapa cara yang dapat digunakan untuk mencocokkan model. Masing-masing cara tersebut memiliki keunggulan dan kelemahan. Dalam penelitian ini, cara yang akan digunakan untuk mencocokkan model adalah dengan cara statistika.

Apabila butir-butir tersebut telah cocok dengan model, berarti syarat-syarat unidimensi, independensi lokal, dan invarian parameter telah terpenuhi untuk masing-masing butir. Dengan demikian, hanya butir-butir yang cocok dengan model yang dapat dianalisis menggunakan *IRT*.

g. Fungsi Informasi

Fungsi informasi adalah informasi tentang hubungan di antara parameter kemampuan responden dengan parameter butir. Hubungan itu berbentuk fungsi dengan berbagai pasangan nilai di antara parameter kemampuan dan butir sehingga disebut fungsi informasi. Dalam teori responsi butir, nilai fungsi informasi berbanding terbalik dengan ketidakpastian. Hal ini berarti bahwa semakin kecil ketidakpastian maka semakin tinggi nilai fungsi informasinya, dan sebaliknya. Fungsi informasi dibedakan menjadi dua jenis, yaitu: 1) fungsi informasi butir, dan 2) fungsi informasi tes.⁴⁴

Fungsi informasi butir (*Item Information Function/IIF*). Fungsi informasi butir merupakan fungsi dari parameter kemampuan θ pada suatu butir, misalnya pada butir ke- i dilambangkan dengan $I_i(\theta)$. Melalui fungsi informasi butir dapat dijelaskan kekuatan suatu butir dan menyatakan kontribusi butir untuk mengungkapkan kemampuan laten (*latent traits*) yang diukur menggunakan butir soal tersebut. Terdapat dua kriteria yang digunakan dalam penentuan fungsi informasi butir. Kriteria pertama adalah perubahan. Pada kriteria ini fungsi informasi butir adalah perubahan yang terjadi pada $P(\theta)$ ketika θ berubah. Semakin besar perubahan $P(\theta)$ ketika θ berubah, semakin tinggi nilai fungsi informasinya dan sebaliknya.

⁴⁴ Dali. S. Naga, *op. cit.*, h. 554.

Fungsi informasi dapat ditemukan melalui kecuraman lengkungan. Semakin curam lengkungan, semakin besar nilai fungsi informasinya. Kecuraman pada lengkungan karakteristik butir adalah derivasi probabilitas jawaban betul terhadap parameter kemampuan. Secara matematis dapat dinyatakan sebagai:

$$Kecuraman = \frac{dP_i(\theta)}{d\theta} \dots\dots\dots (2.23)$$

Kriteria kedua adalah penyebaran. Semakin besar penyebaran, semakin kecil ketepatan sehingga nilai fungsi informasi semakin rendah, dan sebaliknya. Dengan demikian fungsi informasi berbanding terbalik dengan penyebaran atau simpangan baku. Semakin kecil nilai simpangan baku, semakin nesar nilai fungsi informasinya. Secara matematis, simpangan baku dapat dinyatakan sebagai berikut.

$$Simpangan Baku = \sqrt{P_i(\theta) \cdot Q_i(\theta)} \dots\dots\dots (2.24)$$

Berdasarkan dua kriteria di atas, dapat dinyatakan bahwa fungsi informasi berbanding lurus dengan kecuraman dan berbanding terbalik dengan simpangan baku. Secara matematis bentuk fungsi informasi butir adalah sebagai berikut.

$$I_i(\theta) = \frac{\left[\frac{dP_i(\theta)}{d\theta}\right]^2}{P_i(\theta) \cdot Q_i(\theta)} \dots\dots\dots (2.25)$$

Fungsi informasi butir merupakan fungsi dari parameter θ , pada umumnya berbentuk fungsi nonlinear dan memiliki nilai maksimum yang disebut fungsi informasi butir maksimum. Pada fungsi informasi butir maksimum terdapat dua besaran yaitu: (1) letak, dan (2) nilai fungsi informasi butir maksimum. Fungsi informasi butir model L3P dapat dinyatakan sebagai:

$$I(\theta, u_i) = \frac{D^2 a^2 Q_i [(P_i - c_i)^2]}{P_i [(1 - c_i)^2]} \dots\dots\dots (2.26)$$

mencapai maksimum pada:

$$\theta_{max} = b_i + \frac{1}{Da_i} \left\{ \ln \frac{1 + (1 + 8c_i)^{1/2}}{2} \right\} \dots\dots\dots (2.27)$$

sehingga:

$$I(\theta, u_i)_{max} = \frac{D^2 a_i^2}{8(1 - c_i^2)} [1 - 20c_i - 8c_i^2 + (1 + 8c_i)^{3/2}] \dots\dots\dots (2.28)$$

Pada model L2P, maka parameter $c=0$ sehingga rumus (2.28) menjadi:

$$I(\theta, u_i)_{max} = \frac{1}{4} D^2 a^2 = 0,7225 a^2 \dots\dots\dots (2.29)$$

Pada model L1P, maka parameter $a=1$ dan $c=0$ sehingga rumus (2.28) di atas menjadi:

$$I(\theta, u_i)_{max} = \frac{1}{4} D^2 = 0,7225 \dots\dots\dots (2.30)$$

Letak nilai fungsi informasi butir maksimum pada model L1P dan L2P pada parameter taraf sukar butir $\theta_{max} = b_i$, dapat diperoleh dari rumus (2.27).

Sedangkan pada model L3P, letak θ_{max} itu bergeser ke atas b_i dengan jarak pergeseran tergantung kepada nilai C_i .

Fungsi informasi tes (*Test Information Function/TIF*). Tes terdiri atas sejumlah butir, dengan demikian fungsi informasi tes merupakan jumlah dari fungsi informasi butir-butir soal yang ada pada tes tersebut. Akibatnya, nilai fungsi informasi tes akan tinggi bila butir-butir soal penyusun tes memiliki nilai fungsi informasi yang tinggi pula. Semakin tinggi nilai fungsi informasi tes, maka semakin baik mutu tes untuk mengungkapkan kemampuan laten responden.⁴⁵ Pada tes dengan n butir maka fungsi informasi tes, secara matematis dapat dinyatakan sebagai berikut.

$$I(\theta) = \sum_{i=1}^n I_i(\theta) = \sum_{i=1}^n \frac{1}{P_i(\theta) \cdot Q_i(\theta)} \left[\frac{dP_i(\theta)}{d\theta} \right]^2 \dots\dots\dots (2.31)$$

2. Metode Penyetaraan (*Linking Method*)

Dalam teori responsi butir sering dibutuhkan penyamaan skala untuk mencapai skala umum pada estimasi parameter butir yang diperoleh dari kelompok yang berbeda. *Linking* merupakan suatu proses penggabungan skala melalui transformasi dari satu metrik ke metrik lain, sehingga data-data respons dari beberapa kelompok berada dalam satu metrik yang sama.⁴⁶

⁴⁵ Hambleton, R. K., Swaminathan, H., *op. cit.*, h. 104.

⁴⁶ Won-Chan Lee & Jae-Chun Ban, *Comparison of Three IRT Linking Procedures in the Random Groups Equating Design* (Iowa: CASMA, 2007), hh. 1-2.

Kim & Cohen (1998) mendefinisikan *linking* sebagai penyamaan skala dalam *IRT*, yaitu mentransformasikan parameter butir dari metrik yang satu ke metrik yang lain. Data respons kedua kelompok diestimasi secara terpisah sehingga diperoleh data estimasi dari kedua kelompok tersebut. Selanjutnya parameter butir tersebut ditransformasikan dari satu kelompok ke kelompok lainnya.⁴⁷

Sejalan dengan pendapat di atas, Naga (1992) menyatakan bahwa apabila akan membandingkan atau menggabungkan skor hasil tes dari sumber yang berbeda, maka harus dilakukan penyamaan ukuran atau skalanya. Teori dan teknik untuk menyamakan skala yang berbeda dikenal sebagai penyamaan skala atau penskalaan. Setelah skor-skor tersebut disamakan skalanya, barulah skor tersebut dapat digabungkan atau dibandingkan.⁴⁸

Dalam psikometrika proses penyamaan skala harus dilakukan apabila data yang berasal dari kelompok berbeda akan diperbandingkan. Taehoon Kang & Nancy Petersen (2009) menyebutkan bahwa proses penyamaan skala melalui suatu transformasi metrik ke metrik lain disebut dengan

⁴⁷ Allan S. Cohen & Seok-Ho Kim, *An Investigation of Linking Method Under The Graded Response Model*. *Journal Applied Psychological Measurement*, Vol. 22, No. 2, Juni 1998, h. 117.

⁴⁸ Dali S. Naga, *op. cit.*, 343.

linking.⁴⁹ Lebih jauh Seonghoon Kim & Won-Chan Lee (2004) menyatakan bahwa untuk mendapatkan skala umum data yang berasal dari kelompok berbeda, salah satu skala harus dikaitkan dengan yang lain. Dalam hal ini, parameter butir dan parameter kemampuan pada satu skala ditransformasikan ke skala yang lain dengan fungsi linear, proses disebut di sini sebagai *linking*.⁵⁰

Dari uraian dan pendapat-pendapat para ahli di atas dapat disimpulkan bahwa *linking* merupakan suatu proses penyamaan skala data yang bersumber dari kelompok yang berbeda melalui transformasi metrik satu ke metrik yang lainnya, sehingga data-data tersebut berada pada metrik yang sama.

Steven J. Osterlind (1983) mengemukakan bahwa setelah estimasi parameter masing-masing kedua kelompok telah diperoleh, langkah berikutnya yang dilakukan sebelum mendeteksi bias butir adalah menyetarakan skala dan membandingkan kurva karakteristik butir. Ada kemungkinan terdapat perbedaan parameter butir bila diestimasi menggunakan sampel yang berbeda. Meskipun kedua kelompok pada dasarnya pada skala yang sama, keduanya dibedakan oleh transformasi

⁴⁹ Taehoon Kang & Nancy Petersen, *Linking Item Parameters to a Base Scale* (Iowa: ACT, Inc, 2009), hh. 1-2.

⁵⁰ Seonghoon Kim & Won-Chan Lee, *IRT Scale Linking Methods for Mixed-Format Tests* (Iowa: ACT, Inc, 2004), h. 1.

linear. Hal ini terjadi karena pada saat menentukan parameter θ berdasarkan rerata $\mu = 0$ dan standar deviasi $\sigma = 1$ untuk masing-masing kelompok. Dengan kata lain, parameter a_i dan b_i invarian antarkelompok, tetapi kedua parameter tersebut tidak invarian ketika penetapan θ awal bebas dipertukarkan pada setiap proses parameterisasi.⁵¹

Lebih lanjut Osterlind (1983) menyatakan bahwa bila skala θ telah disetarakan, maka membandingkan kurva karakteristik butir kedua kelompok sudah dapat dilakukan. Untuk menyetarakan parameter kedua kelompok dapat dilakukan dengan rumus transformasi linear sebagai berikut.

$$a_{i2}^* = \frac{a_{i2}}{A} \dots\dots\dots (2.32)$$

$$b_{i2}^* = Ab_{i2} + K \dots\dots\dots (2.33)$$

$$c_{i2}^* = c_{i2} \dots\dots\dots (2.34)$$

$$A = \frac{\sigma_{bi1}}{\sigma_{bi2}} \dots\dots\dots (2.35)$$

$$K = \mu_{bi1} - A\mu_{bi2} \dots\dots\dots (2.36)$$

Dengan μ_{bi1} merupakan rerata taraf sukar butir ke-i kelompok ke-1 (referensi) dan σ_{bi2} merupakan simpangan baku taraf sukar butir ke-i kelompok ke-2 (fokal). Tanda (*) menyatakan nilai transformasi sehingga estimasi parameter taraf sukar butir ke-i kelompok kedua (fokal) adalah b_{i2}^* , daya pembeda butir

⁵¹ Osterlind, S., J., *Test Item Bias* (USA: Sage Publication, 1983), hh. 65-66.

ke-i kelompok kedua (fokal) adalah a_{i2}^* , dan kebetulan menjawab betul adalah c_{i2}^* dari kelompok kedua (fokal) yang skalanya disamakan dengan kelompok pertama (referensi). Sedangkan parameter kemampuan peserta kelompok kedua (fokal) pada butir ke-j disamakan skalanya dengan kelompok pertama dengan rumus transformasi linear sebagai berikut.

$$\theta_{j2}^* = A\theta_{j2} + K \dots\dots\dots (2.37)$$

Sejalan dengan pendapat Osterlind, ahli lainnya Camilli & Shepard (1994) juga menyatakan bahwa sebelum dilakukan pendeteksian bias butir menggunakan metode *DIF*, parameter butir harus disamakan atau parameter butir skala kelompok fokal dan kelompok referensi harus berada dalam skala yang sama.⁵² Dalam penelitian ini metode penyetaraan parameter yang digunakan adalah metode rerata dan sigma. Metode rerata dan sigma menggunakan rerata dan simpangan baku dari estimasi taraf sukar butir dari kedua kelompok peserta tes. Metode ini sangat cocok digunakan untuk data model dikotomus.

Peter Baldwin (2013) menyatakan bahwa metode rerata dan sigma merupakan metode penyetaraan parameter yang sangat familiar dan memiliki akurasi yang tinggi. Perhitungannya sangat sederhana, tidak menggunakan rumus-rumus yang rumit. Transformasi linear ini memuat bentuk *slope* and

⁵² Camilli & Shepard, *op. cit.*, h. 62.

intercept. Oleh karena itu, metode penyetaraan rerata dan sigma sangat populer digunakan dalam penyetaraan parameter.⁵³

Lebih lanjut Seonghoon Kim & Won-Chan Lee (2006) menyebutkan bahwa metode rerata dan sigma lebih dikenal sebagai *moment methods*, yang memiliki ciri perhitungan yang sangat sederhana, namun memiliki akurasi yang tinggi. Metode rerata dan sigma Estimasi menghasilkan koefisien penyetaraan sangat akurat. Keunggulan-keunggulan tersebut sering dijadikan alasan untuk memilih metode rerata dan sigma untuk penyetaraan parameter.⁵⁴

Menurut Won-Chan Lee & Jae-Chun Ban (2010), metode rerata dan sigma sebagai salah satu metode penyetaraan terpisah, menggunakan momen estimasi parameter butir untuk mendapatkan skala transformasi parameter. Berbeda dengan metode Haebara dan Stocking Lord, estimasi skala transformasi dihitung menggunakan kurva karakteristik butir masing-masing kelompok. Perhitungan skala penyetaraan parameter menggunakan metode rerata dan sigma lebih sederhana dibandingkan dengan metode Haebara dan Stocking Lord.⁵⁵

⁵³ Peter Baldwin, On Mean-Sigma Estimators and Bias. *British Journal of Mathematical and Statistical Psychology* (2013), 66, 277–289 (diakses 4 Maret 2016), h. 278.

⁵⁴ Seonghoon Kim & Won-Chan Lee, An Extension of Four IRT Linking Methods for Mixed-Format Tests. *Journal of Educational Measurement Spring 2006, Vol. 43, No. 1, pp. 53–76* (diakses 5 Maret 2016), h. 59.

⁵⁵ Won-Chan Lee & Jae-Chun Ban, A Comparison of IRT Linking Procedures. *Journal Applied Measurement In Education*, 23: 23–48, 2010 (diakses 5 Maret 2016), h. 27.

Walaupun metode rerata dan sigma sangat sederhana dalam perhitungan dan sangat familiar, di sisi lain metode rerata dan sigma juga memiliki kelemahan sebagaimana diungkapkan oleh Deping Li, Yanlin Jiang, dan Alina A. von Davier (2012). Metode rerata dan sigma (*moment methods*) merupakan metode penyetaraan terpisah (*sparated linking methods*), memiliki peluang kesalahan yang lebih besar daripada metode penyetaraan serentak (*concurrent calibration*) seperti metode kurva karakteristik (*characteristic curve methods*).⁵⁶

Langkah-langkah yang dilakukan untuk penyamaan skala menggunakan metode rerata dan sigma adalah sebagai berikut.

- a) Mengestimasi parameter taraf sukar butir dari kelompok fokal b_F dan kelompok referensi yaitu b_R secara terpisah. Parameter ini memiliki rerata dan simpangan baku. Rerata dan simpangan baku kelompok fokal berturut-turut dilambangkan dengan μ_{b_F} dan σ_{b_F} . Sedangkan rerata dan simpangan baku kelompok referensi berturut-turut dilambangkan dengan μ_{b_R} dan σ_{b_R} .
- b) Mentransformasikan taraf sukar butir ke- j dari kelompok fokal ke kelompok referensi menggunakan rumus (2.32) yaitu $b_{jF}^* = Ab_{jF} + K$, dengan b_{jF} merupakan taraf sukar butir ke- j pada kelompok fokal sebelum dilakukan

⁵⁶ Deping Li, Yanlin Jiang, dan Alina A. von Davier, The Accuracy and Consistency of a Series of IRT True Score Equatings. *Journal of Educational Measurement Summer 2012, Vol. 49, No. 2, pp. 167–189* (diakses 5 Maret 2016), h. 168.

penyamaan skala, sedangkan b_{jF}^* merupakan taraf sukar butir ke-j pada kelompok fokal setelah dilakukan penyamaan skala. Nilai b_{jF}^* pada kelompok fokal berada pada skala yang sama dengan b_{jR} pada kelompok referensi. Koefisien A dan K dihitung menggunakan rumus (2.35) dan (2.36) adalah: $A = \frac{\sigma_{bjR}}{\sigma_{bjF}}$ dan $K = \mu_{bR} - A\mu_{bF}$.

- c) Mentransformasikan daya pembeda dan kebetulan menjawab betul pada butir ke-j dari kelompok fokal ke kelompok referensi, dengan rumus (2.33) dan (2.34) yaitu: $a_{jF}^* = \frac{a_{jF}}{A}$ dan $c_{jF}^* = c_{jF}$. Dengan a_{jF} merupakan daya pembeda butir ke-j kelompok fokal sebelum ditransformasi (penyamaan skala). Sedangkan a_{jF}^* merupakan daya pembeda butir ke-j kelompok fokal setelah dilakukan penyamaan skala. Nilai a_{jF}^* pada kelompok fokal berada pada skala yang sama dengan a_{jF} pada kelompok referensi. Sedangkan nilai c_{jF}^* merupakan nilai kebetulan menjawab betul butir ke-j pada kelompok fokal setelah dilakukan penyamaan skala. Nilai c_{jF}^* sama dengan nilai c_{jF} sebelum dilakukan penyamaan skala.

3. Bias Butir

Pada pembahasan berikut ini, diuraikan landasan teori yang terkait dengan bias butir. Pembahasan dimulai dari pengertian bias butir, jenis-jenis bias butir, bias gender, metode *DIF* yang diperbandingkan, dan perbandingan sensitivitas metode *DIF*.

a. Pengertian

Camilli & Shepard (1994) mengemukakan bahwa terdapat perbedaan pengertian antara *Differential Item Functioning (DIF)* dan bias butir. Namun banyak penulis menganggap kedua istilah itu memiliki pengertian yang sama dan dapat dipertukarkan. *DIF* diartikan sebagai pengukuran taraf sukar relatif butir yang menyimpang secara berlebihan pada kelompok berbeda dan responden yang memiliki kemampuan sama. Sedangkan bias butir diartikan sebagai ketidakvalidan suatu butir, sehingga butir tersebut tidak memberikan peluang yang sama menjawab benar pada dua kelompok yang berbeda pada peserta dengan kemampuan yang sama. Idealnya statistik *DIF* itu digunakan untuk mencari butir tes yang berfungsi secara berbeda untuk kelompok yang berbeda kemudian setelah dilakukan analisis logis mengapa butir-butir itu secara relatif dirasakan lebih sulit, sekelompok butir bias kemudian dapat diidentifikasi sebagai bias dan tentu saja harus dikeluarkan dari tes. Namun, ketika tidak menganalisis penyebab timbulnya bias maka kedua istilah itu dapat digunakan secara bergantian. Dengan demikian *DIF* merupakan metode statistika yang digunakan untuk mendeteksi adanya bias butir.⁵⁷

Mohammad Salehi dan Alireza Tayebi (2012) menyatakan bahwa *DIF* adalah suatu metode dalam ilmu Psikometrika yang digunakan untuk

⁵⁷ Camilli & Shepard, *Methods for Identifying Biased Test Items* (New Delhi: Sage Publication Inc, 1994), h. 16.

mendeteksi bias butir. Metode *DIF* digunakan untuk menentukan apakah sebuah tes layak diperbandingkan jika digunakan untuk mengukur parameter kemampuan responden, yang berasal dari kelompok-kelompok yang berbeda.⁵⁸

Badrun Kartowagiran (2005) mengemukakan bahwa butir yang bias adalah butir yang tidak memberi peluang yang sama untuk menjawab benar pada peserta tes yang memiliki kemampuan sama hanya karena berasal dari kelompok yang berbeda. Hal ini berarti bahwa bias butir terjadi manakala individu-individu dengan kemampuan sama dari kelompok berbeda memiliki peluang berbeda untuk menjawab benar butir yang sama.⁵⁹

Menurut Likun Hou, Jimmy de la Torre, dan Ratna Nandakumar (2014), prosedur tradisional untuk mendeteksi *DIF* berdasarkan model IRT merupakan teknik menginvestigasi peserta tes dengan level kemampuan yang sama, namun berasal dari kelompok yang berbeda mengakibatkan adanya perbedaan probabilitas menjawab benar suatu butir tertentu. Hal itu berarti bahwa adanya *DIF* pada suatu butir tes dapat mengakibatkan estimasi terhadap parameter butir dan kemampuan responden dapat berubah.⁶⁰

⁵⁸ Mohammad Salehi dan Alireza Tayebi, Differential Item Functioning: Implications for Test Validation. *Journal of Language Teaching and Research*, Vol. 3, No. 1, pp. 84-92, January 2012 ISSN 1798-4769 (diakses 5 Maret 2016), h. 87.

⁵⁹ Badrun Kartowagiran, Perbandingan Berbagai Metode untuk Mendeteksi Bias Butir. *Disertasi*. (Yogyakarta: UGM, 2005), h. 17.

⁶⁰ Likun Hou, Jimmy de la Torre, dan Ratna Nandakumar, Differential Item Functioning Assessment in Cognitive Diagnostic Modeling: Application of theWald Test to Investigate

Selanjutnya Osterlind (1983) menyatakan bahwa pengertian bias butir pada pengukuran merupakan suatu istilah yang mempunyai makna rasial, menekan atau terlalu fanatik dengan objek yang akan diukur. Bias butir merupakan suatu kondisi butir yang tidak adil (*unfair*), tidak konsisten, terkontaminasi oleh faktor-faktor di luar aspek yang hendak diukur, dan kesalahan penggunaan tes. Hal ini berarti bahwa bias dalam butir dan pengukuran mempunyai makna yang tidak mendukung sifat tes yang konsisten dan sah.⁶¹

Di bagian lain Osterlind (1983) menjelaskan bahwa pembicaraan mengenai bias dalam tes psikologi berhubungan dengan validitas konstruk bagi butir tes. Sejauh mana butir tes (atau perangkat butir) dapat mengukur suatu konstruk atau sifat yang jelas dan tunggal. Apabila butir tes memiliki validitas konstruk yang sama untuk semua orang dalam satu populasi maka peserta tes dengan kemampuan sama haruslah memiliki kesempatan yang sama untuk menjawab benar.

Sementara itu Mazor (1995) mengemukakan bahwa suatu tes dikatakan bias jika dua orang peserta tes dengan kemampuan yang sama dari kelompok yang berbeda, tidak memperoleh peluang menjawab benar

DIF in the DINA Model. *Journal of Educational Measurement Spring 2014, Vol. 51, No. 1, pp. 98–125* (diakses 5 Maret 2016), h. 101.

⁶¹ Osterlind, S. J., *op. cit.*, h. 10.

yang sama.⁶² Dengan demikian butir tes yang tidak bias adalah butir tes yang diharapkan akan memberikan peluang menjawab benar yang sama pada peserta tes yang memiliki kemampuan yang sama dari kelompok yang berbeda.

Ada dua pendekatan dalam memaknai istilah bias. Pertama, bias dapat diketahui dengan melihat kegunaan skor tes yang akan diperoleh. Dalam hal ini kegunaannya dapat disesuaikan dengan tujuannya, misalnya seleksi masuk sekolah, seleksi penerimaan karyawan baru, dan berbagai macam program dan pelayanan yang berdasarkan pada suatu tingkat skor tes yang tinggi. Pendekatan ini memerlukan perbandingan hasil tes suatu kelompok dengan beberapa kelompok lainnya, sehingga akan terdapat sejumlah bentuk bias dalam seleksi. Pendekatan ini merupakan bias eksternal dalam butir tes. Kedua, adanya dugaan butir itu bias. Pendekatan ini dilakukan dengan melihat butir-butir tes itu sendiri, apakah ada yang menyalahi kriteria-kriteria yang telah ditentukan. Melegitimasi penggunaan suatu tes agar seluruh peserta memperoleh butir-butir tes yang adil sesuai dengan prestasi atau kemampuannya masing-masing, merupakan suatu permasalahan yang sangat krusial. Oleh karena itu, tes harus dibuat sedemikian sehingga adanya perbedaan skor tes di antara kelompok-

⁶² Mazor, K. M., Using Logistic Regression and Mantel Haenszel With Multiple Ability Estimates to Detect Differential Item Functioning. *Journal of Education Measurement*. 32 (2), h. 133.

kelompok peserta tes, murni disebabkan karena perbedaan dari pemahaman isi tes yang diukur. Pendekatan ini disebut pendekatan internal dalam butir.⁶³

Hulin, C. L., Drasgow, F., Parsons, C. K. (1983) mengemukakan bahwa bias butir terjadi ketika individu dengan kemampuan yang sama tetapi berasal dari subpopulasi yang berbeda memiliki probabilitas yang berbeda menanggapi suatu butir dengan benar. Bias butir dapat disebabkan oleh penggunaan hal yang sangat spesifik seperti pengetahuan tentang subkultur yang tidak secara luas dibagi dalam populasi, butir yang mengacu pada konsep tidak relevan beberapa subpopulasi, atau butir telah diterjemahkan kurang baik ke dalam bahasa asing lainnya.⁶⁴

Hambleton & Swaminathan (1985) mendefinisikan bahwa bias butir dan bias tes menunjukkan adanya berat sebelah terhadap kelompok-kelompok yang berbeda, sekalipun peserta tes memiliki kemampuan yang sama. Artinya, ada kelompok peserta yang diuntungkan, di sisi lain ada kelompok yang dirugikan. Dalam pengertian tersebut, permasalahan yang dijumpai cukup jelas yaitu adanya variabel lain yang berkontribusi terhadap adanya perbedaan tersebut.⁶⁵

⁶³ Osterlind, S. J., *op. cit.*, h. 11.

⁶⁴ Hulin, C. L., Drasgow, F., Parsons, C. K., *Item Response Theory. Application to Psychological Measurement* (USA: Dow Jones-Irwin, 1983), h. 152.

⁶⁵ Hambleton, R. K., Swaminathan, H., *op. cit.*, hh. 281-282.

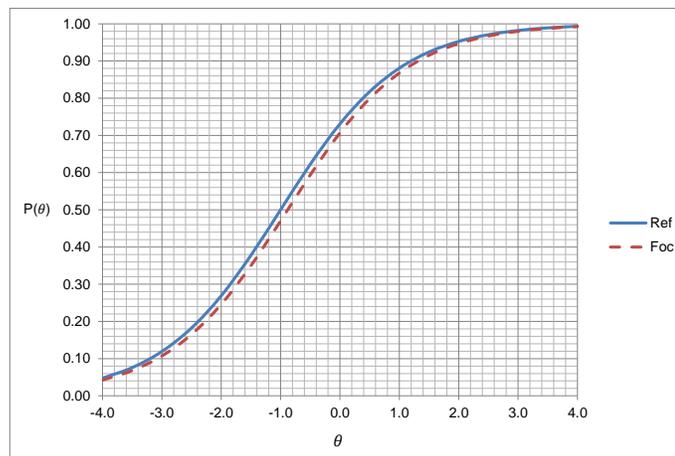
Hambleton, Swaminathan dan Roger (1991) menjelaskan pengertian *DIF* yang dikaitkan dengan kurva karakteristik butir. Bila terdapat dua kelompok peserta tes dengan kemampuan yang sama menjawab sejumlah butir, selanjutnya dilakukan estimasi terhadap parameter butir dari kedua kelompok yang berbeda, lalu digambarkan kurva karakteristik butir kedua kelompok. Apabila luas daerah di antara kedua karakteristik butir tidak sama dengan nol atau selisihnya cukup besar dikatakan butir tersebut mengandung bias butir. Sebaliknya, bila luas daerah di antara kedua karakteristik butir sama dengan nol (berimpit) atau selisihnya sangat kecil dikatakan butir tersebut tidak bias.⁶⁶

Lebih lanjut Hambleton R. K, Swaminathan H., dan Roger, H. J., (1991) mengemukakan bahwa definisi *DIF* yang dapat diterima oleh para ahli psikometrika adalah butir menunjukkan *DIF* jika individu memiliki kemampuan yang sama, tetapi berasal dari kelompok berbeda, tidak memiliki probabilitas yang sama untuk menjawab butir itu dengan tepat. Dengan kata lain, butir menunjukkan *DIF* jika fungsi respon butir seluruh subkelompok yang berbeda tidak identik. Sebaliknya, butir tidak menunjukkan *DIF* jika fungsi karakteristik butir seluruh subkelompok yang berbeda adalah identik.

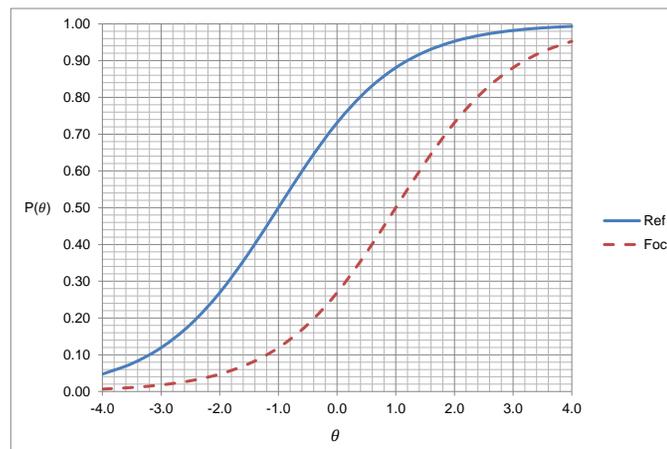
Dengan demikian, butir-butir yang bias mengakibatkan probabilitas responden dengan kemampuan yang sama untuk menjawab benar butir

⁶⁶ Hambleton, R. K., Swaminathan, H., dan Roger, H. J., *Fundamentals of Item Response Theory* (London: Sage Publication Inc, 1991), h. 110.

tertentu tidak sama hanya karena berasal dari kelompok yang berbeda. Apabila probabilitas menjawab benar kedua kelompok digambarkan dalam kurva karakteristik butir, akibatnya kedua kurva karakteristik butir tidak identik. Sebaliknya, apabila kedua kurva karakteristik butir identik, menyatakan bahwa butir tersebut tidak bias. Gambar 2.6 dan gambar 2.7 menyajikan ilustrasi kurva karakteristik butir yang tidak bias dan bias.



Gambar 2.6 Kurva Karakteristik Butir Tidak Bias



Gambar 2.7 Kurva Karakteristik Butir Bias

Sejalan dengan pendapat para ahli di atas, Naga (2013) menyatakan bahwa butir dikatakan bias jika dua kelompok responden berkemampuan sama memperoleh skor berbeda pada butir itu. Dalam hal ini ada kelompok yang merasa dirugikan oleh butir yang bias karena memperoleh skor yang lebih rendah dari skor kelompok lainnya yang berkemampuan sama.⁶⁷

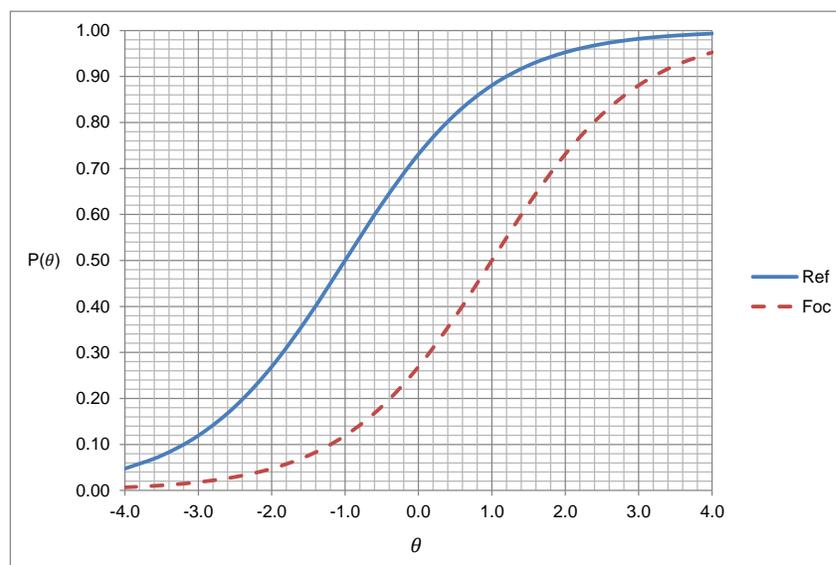
Berdasarkan pendapat-pendapat para ahli di atas, maka dapat disimpulkan bahwa yang dimaksud dengan bias butir adalah butir yang mengakibatkan adanya skor tes yang berbeda bila diujikan pada peserta dengan kemampuan yang sama, namun mereka berasal dari kelompok yang berbeda. Hal itu terjadi karena adanya variabel lain di luar sasaran pengukuran yang ikut berkontribusi terhadap terjadinya perbedaan peluang menjawab benar pada peserta dengan kemampuan sama yang berasal dari kelompok yang berbeda.

b. Jenis-Jenis Bias Butir

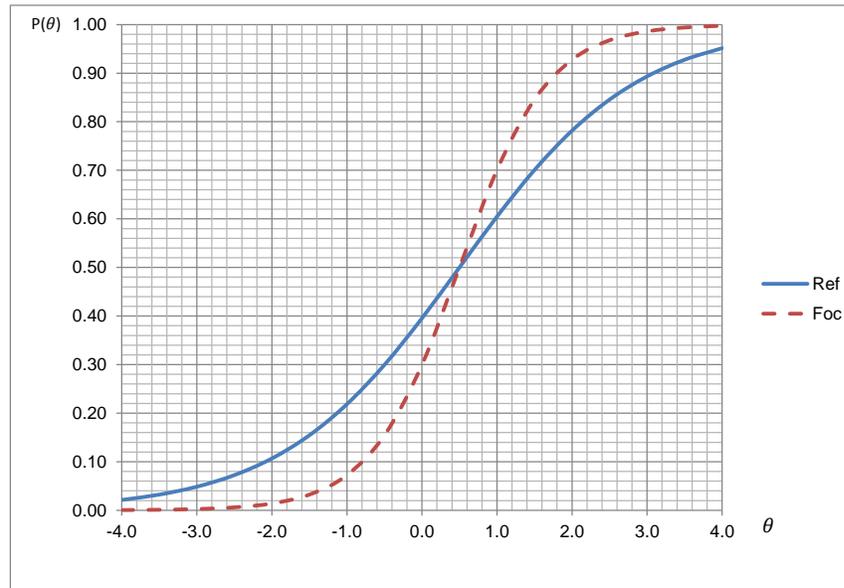
Butir-butir yang dikategorikan bias dapat menguntungkan salah satu kelompok untuk semua interval kemampuan. Pada konsidi lainnya, bias butir juga dapat menguntungkan peserta tes pada kelompok tertentu hanya pada level kemampuan tertentu saja. Misalnya pada interval kemampuan rendah menguntungkan kelompok A, sedangkan pada interval kemampuan tinggi menguntungkan kelompok B. Sesuai dengan pengertian bias butir yang telah

⁶⁷ Dali S. Naga, *op. cit.*, hh. 607-608.

dipaparkan di atas, terdapat dua jenis bias butir, yaitu bias butir *uniform* (konsisten) dan bias butir *non-uniform* (tidak konsisten). Finch & French (2007) mengemukakan bahwa bias butir *uniform* muncul jika keuntungan salah satu kelompok terhadap kelompok lainnya terjadi pada setiap level kemampuan, sedangkan bias butir *non-uniform* muncul jika keuntungan salah satu kelompok terhadap kelompok lainnya tidak terjadi pada setiap level kemampuan. Dilihat dari kurva karakteristik butir, pada bias butir *uniform* (konsisten) grafik kelompok yang satu selalu berada di bawah kelompok yang lainnya. Sedangkan pada bias butir *non-uniform*, grafik kelompok yang satu tidak selalu berada di bawah kelompok yang lainnya dengan kata lain kedua kurva saling berpotongan. Untuk lebih jelasnya, perhatikan gambar 2.8 dan 2.9 di bawah ini.



Gambar 2.8 Bias Butir *Uniform*



Gambar 2.9 Bias Butir *Non-Uniform*

c. Bias Gender

World Health Organization, WHO (2009) merumuskan konsep dasar perbedaan antara jenis kelamin dan gender. Pengertian gender merujuk pada perbedaan peran dan tanggung jawab antara perempuan dan laki-laki sejak lahir, tumbuh kembang dan besar melalui proses sosialisasi di lingkungan keluarga dan masyarakat. Sedangkan jenis kelamin mengacu pada perbedaan karakteristik biologis dan fisiologis yang membedakan perempuan dan laki-laki. Jenis kelamin bersifat kodrati dan universal (berlaku di mana saja) dan tidak bisa dipertukarkan satu sama lain.⁶⁸

⁶⁸ World Health Organization, *Integrating gender into HIV/AIDS programmes in the health sector Tool to improve responsiveness to women's needs* (Geneva: WHO Press, 2009), h.

Adanya perbedaan peran gender melalui pemisahan kepantasan untuk perempuan dan kepantasan untuk laki-laki, sangat dipengaruhi oleh faktor-faktor lingkungan sosial. Perbedaan peran gender tidak bersifat universal, tetapi dapat berbeda antara satu kebudayaan dengan kebudayaan lainnya, antara etnis satu dengan etnis yang lainnya, dan dapat berubah seiring dengan perkembangan zaman. Gender mengacu pada peran, perilaku, kegiatan serta karakteristik sosial lainnya yang dibentuk oleh suatu masyarakat atau budaya tertentu berdasarkan persepsi yang pantas untuk perempuan atau pantas untuk laki-laki.

Dalam konsep jenis kelamin (*sex roles concept*), secara kodrat struktur anatomi wanita memungkinkan mereka untuk melaksanakan fungsi reproduksi seperti mengandung, melahirkan, menyusui dan menstruasi yang tidak dapat dilakukan oleh lelaki. Berdasarkan teori rekonstruksi sosial (*social reconstruction*) fungsi reproduksi diperluas ke tugas-tugas domestik di dalam rumah tangga, seperti membesarkan anak, mencuci, dan memasak. Konstruksi sosial ini telah menimbulkan diskriminatif terhadap perempuan yang menyebabkan mereka kurang berdaya dalam domain publik. Akibatnya perempuan menjadi segmen masyarakat yang secara ekonomis kurang produktif. Lebih jauh akan menyebabkan kualitas hidup relatif lebih rendah yang akan menyebabkan rendahnya kualitas anak dan masa depan masyarakat. Salah satu gejala adalah pembakuan peran jenis kelamin

(*gender stereotype*); laki-laki ditempatkan sebagai peran fungsi publik dan produktif, sedangkan perempuan pemeran domestik dan reproduktif. Dalam banyak kasus, perlakuan tidak adil banyak menimpa perempuan di rumah, di tempat kerja, maupun di tengah-tengah masyarakat. Ketidakadilan itu tercipta karena dikonstruksi oleh budaya dan lembaga atau dilembagakan. Dampak dari semua dapat menyebabkan terhambatnya kemajuan bangsa.⁶⁹

Ulric Neisser et al. (1996) menyatakan adanya perbedaan kemampuan dasar seseorang (intelegensi) juga dipengaruhi oleh faktor-faktor sosial dan biologis. Pada faktor-faktor sosial, kondisi-kondisi yang dapat mempengaruhi intelegensi antara lain perbedaan lingkungan, perbedaan gender, lingkungan keluarga, kebiasaan-kebiasaan yang berlaku setempat, dan kebudayaan. Sedangkan faktor-faktor biologis terkait dengan jenis kelamin laki-laki dan perempuan, genetika, serta fungsi hormon dalam tubuh.⁷⁰

Lebih lanjut Setia Iriyanto dan Eny Winaryati (2010) dalam penelitiannya menyatakan bahwa tingkat pendidikan guru laki-laki relatif lebih tinggi dibandingkan dengan guru perempuan. Hal ini disebabkan karena kesempatan untuk kuliah bagi guru laki-laki lebih longgar dibandingkan

⁶⁹ Setia Iriyanto dan Eny Winaryati, Perbedaan Persepsi Antar Jenis Kelamin terhadap Peran Gender Dalam Keluarga Dan Masyarakat : Antara Harapan dan Kenyataan pada Guru-guru SD di Wilayah Kecamatan Tembalang Kota Semarang. *Prosiding Seminar Nasional Unimus*, ISBN 978.979.704.883.9, 2010, hh. 210-211.

⁷⁰ Neisser, U. et. al., *Intelligence: Knowns and Unknowns*. American Psychological Association, Inc., 1996, hh. 37-38.

dengan perempuan. Persoalan mendasar yang ada adalah karena kesibukan rumah tangga yang tidak bisa dipisahkan dari kehidupan perempuan.⁷¹

Yeni Tri Asmaningtyas (2012) menyatakan bahwa berkaitan dengan pembelajaran matematika di sekolah yang melibatkan siswa laki-laki dan perempuan, pada umumnya perempuan lebih baik dalam ingatan dan laki-laki lebih baik dalam hal berpikir logis.⁷² Lebih lanjut Kartini Kartono (1989) dalam Yeni Tri Asmaningtyas menyatakan bahwa betapapun baik dan cemerlangnya intelegensi perempuan namun pada intinya perempuan hampir-hampir tidak pernah mempunyai ketertarikan yang menyeluruh pada soal-soal teoretis seperti pada laki-laki, perempuan lebih tertarik pada hal-hal yang praktis dan konkret. Sedangkan laki-laki lebih tertarik pada soal-soal yang teoretis dan abstrak. Sejalan dengan pendapat di atas Benbov dan Stanley (1992) menyatakan bahwa dilihat hubungan jenis kelamin dan hasil belajar matematika, siswa laki-laki lebih unggul dalam matematika terutama dalam soal-soal spasial. Sehingga dalam topik-topik tertentu dalam geometri dan ilmu ukur ruang, hasil belajar siswa laki-laki lebih baik daripada siswa perempuan. Dengan demikian, secara alamiah terdapat keunggulan dan kelemahan antara siswa laki-laki dan perempuan pada topik-topik matematika tertentu.

⁷¹ Setia Iriyanto dan Eny Winaryati, *op. cit.*, h.2012

⁷² Yeni Tri Asmaningtyas, "Kemampuan Matematika Laki-laki dan Perempuan", <http://download.portalgaruda.org/article.php?article=115727&val=5278> (diakses 10 Oktober 2015), hh. 3-4.

Yanmei Li, Allan S. Cohen, dan Robert A. Ibarra (2004) menyelidiki karakteristik butir soal Matematika ditinjau dari perbedaan gender. Bias butir dapat terjadi karena adanya perbedaan gender dan etnik. Perbedaan gender menyebabkan adanya perbedaan strategi dalam penyelesaian soal-soal matematika. Problem matematika dapat diklasifikasikan menjadi problem konvensional dan nonkonvensional. Problem konvensional seperti problem rutin (*textbook*) yang dapat diselesaikan dengan algoritma matematika biasa, sedangkan problem nonkonvensional umumnya memerlukan cara penyelesaian yang tidak umum dilakukan, memerlukan wawasan, serta cara berpikir kritis dan kreatif. Hasil penelitiannya menyatakan bahwa perempuan umumnya lebih unggul dibandingkan dengan laki-laki dalam menyelesaikan masalah konvensional yang mengikuti pola keteraturan (algoritma). Sebaliknya, laki-laki lebih unggul daripada perempuan untuk menyelesaikan masalah nonkonvensional yang menuntut kemampuan berpikir tingkat tinggi (*higher order thinking skills*). Perempuan juga lebih baik dalam hal kecepatan mengakses informasi dan menyelesaikan soal-soal yang memerlukan ingatan, sedangkan laki-laki lebih unggul dalam menyelesaikan problem matematika yang menuntut inovasi dan kreativitas tinggi.⁷³

Dian Prastiwi Darmastini dan Abdul Haris Rosyidi (2014) menyatakan bahwa representasi merupakan ungkapan dari gagasan-gagasan atau ide-ide

⁷³ Yanmei Li, Allan S. Cohen, dan Robert A. Ibarra, Characteristics of Mathematics Items Associated With Gender DIF. *International Journal of Testing*, Vol. 4, No. 2, 2004, Lawrence Erlbaum Associates, Inc. (diakses 11 September 2015), hh. 118-119.

matematika yang ditampilkan siswa dalam upayanya untuk mencari suatu solusi dari masalah yang sedang dihadapinya. Representasi dapat berupa tabel, grafik, foto, gambar, atau persamaan. Representasi sangat berperan dalam proses penyelesaian masalah matematis, sangat memungkinkan bagi siswa untuk mencoba berbagai representasi sebagai wujud dari strategi-strateginya untuk menyelesaikan suatu masalah matematis. Penggunaan representasi yang beragam dalam menyelesaikan suatu soal disebut multi representasi. Faktor yang berpengaruh terhadap kemampuan penyelesaian masalah matematika adalah gender. Hal tersebut memungkinkan terdapat adanya perbedaan strategi-strategi yang digunakan oleh siswa laki-laki dan perempuan dalam cara penyelesaian masalah.⁷⁴

Lebih lanjut dikemukakan bahwa siswa perempuan, secara umum, lebih unggul dalam bidang bahasa dan menulis, sedangkan siswa laki-laki lebih unggul dalam bidang matematika karena kemampuan-kemampuan ruangnya yang lebih baik. Siswa laki-laki dan perempuan cenderung menggunakan representasi persamaan dan tabel, sedangkan perempuan juga cenderung memberikan representasi verbal. Siswa laki-laki cenderung memberikan dua sampai tiga representasi untuk setiap soalnya, sedangkan perempuan cenderung memberikan tiga sampai empat representasi untuk

⁷⁴ Dian Prastiwi Darmastini dan Abdul Haris Rosyidi, *Multi Representasi Siswa SMP dalam Menyelesaikan Soal Terbuka Matematika Ditinjau dari Perbedaan Gender*. Jurnal Ilmiah Pendidikan Matematika MATHEdunesa, Volume 3 Nomor 1 Tahun 2014. <http://ejournal.unesa.ac.id/article/9928/30/article.doc> (diakses 10 Oktober 2015), h. 57.

setiap soalnya. Jadi multi representasi lebih banyak dimunculkan oleh siswa perempuan dari pada laki-laki.

Dari uraian di atas dapat disimpulkan bahwa perbedaan jenis kelamin dan gender dapat mempengaruhi kemampuan seseorang untuk memecahkan suatu masalah. Peserta didik laki-laki dan perempuan memiliki sejumlah perbedaan yang disebabkan oleh adanya perbedaan perlakuan, kebiasaan, perasaan, dan kebudayaan antara etnis satu dengan yang lainnya. Di samping itu, pada umumnya peserta didik laki-laki dilihat dari postur tubuhnya lebih kuat bila dibandingkan dengan peserta didik perempuan. Oleh karena itu, apabila peserta didik laki-laki dan perempuan diberikan menjawab sejumlah butir soal yang sama, dapat terjadi perbedaan peluang menjawab benar antar kedua kelompok tersebut. Kondisi tersebut dapat menyebabkan terjadi bias butir karena perbedaan jenis kelamin.

d. Metode *DIF* yang Diperbandingkan

Untuk mendeteksi bias butir, ada beberapa metode *DIF* yang dapat digunakan. Adams dalam Keeves (1992) menyatakan bahwa metode *DIF* yang dapat digunakan untuk mendeteksi bias butir meliputi: analisis faktor, pengujian daya pembeda butir menggunakan *point-biserial* dan *korelasi parsial*, pengujian tingkat kesukaran butir menggunakan berbagai transformasi, metode Anova, teori responsi butir atau metode *latent traits*,

pendekatan khi-kuadrat (*Chi-square*), model *log-linear* dan metode statistik Mantel Haenszel.⁷⁵

Stark & Chernyshenko (2002), mengemukakan ada beberapa metode *DIF*, yang dapat digunakan untuk mendeteksi bias butir baik dengan statistik parameterik dan non parameterik. Jika menggunakan statistik parameterik, maka dapat digunakan metode khi-kuadrat Lord (*Lord's chi-square*), tes rasio kebolehjadian (*likelihood ratio test*), dan metode luas daerah bertanda dan tidak bertanda (*signed and unsigned area methods*). Sedangkan untuk metode statistik non parameterik, dapat menggunakan metode *SIBTEST* (*The Simultaneous Item Bias Test*) dan *Mantel-Haenszel*. Baik menggunakan pendekatan statistik parameterik dan non parameterik, metode-metode tersebut hanya diterapkan pada tes yang mengukur satu kemampuan saja (unidimensi) saja.⁷⁶

Sementara itu, Camilli & Shepard (1994) mengemukakan terdapat 5 metode *DIF* yang berbasis *IRT* untuk mendeteksi bias butir. Kelima metode yang dimaksudkan adalah: (1) *test of b difference*, (2) *item drift method*, (3) *Lord's Chi-Square*, (4) *empirical sampling distributions for DIF indices*, dan (5)

⁷⁵ Adams, R. J., Item Bias dalam J. P. Keeves, (Ed)., op. cit., h. 180.

⁷⁶ Stark, S., & Chernyshenko, O., Detection of Differential Item/Test Functioning (DIF/DTF) Using IRT. Diakses tanggal 5 Agustus 2015, dari <http://www.work.psych.uiuc.edu/irt/>, h. 4.

model comparison measures (likelihood tes rasio). Kelima metode tersebut memiliki keunggulan dan kelemahan.⁷⁷

Lebih lanjut Camilli dan Shepard (1994) menegaskan kembali bahwa metode *DIF* untuk mendeteksi bias butir yang menggunakan pendekatan teori klasik sebaiknya tidak digunakan lagi. Meskipun metode bias klasik ini secara intuitif menarik, namun secara teknis tidak akurat. Apabila dua kelompok yang dibandingkan memiliki perbedaan dalam penampilan rata-rata, teknik statistik klasik ini akan merancukan antara kesukaran butir dan bias sehingga ditegaskan bahwa berbagai metode untuk mendeteksi bias yang menggunakan pendekatan teori klasik itu kurang tepat sehingga tidak direkomendasikan untuk digunakan. Butir dikatakan bias bila peserta tes berkemampuan sama dari kelompok berbeda memiliki kesempatan yang tidak sama untuk menjawab benar butir yang sama.⁷⁸

Berdasarkan rekomendasi yang disampaikan oleh Camilli dan Shepard (1994) di atas, maka metode *DIF* yang diperbandingkan dalam penelitian ini adalah metode *DIF* yang berdasarkan teori responsi butir. Adapun metode *DIF* yang dibandingkan adalah: 1) Uji Beda Taraf Sukar, 2) Khi-kuadrat Lord, dan 3) Distribusi Sampling Empiris.

⁷⁷ Camilli, G. & Shepard, L. A., *op. cit.*, hh. 71-77.

⁷⁸ Camilli, G. & Shepard, L. A., *op. cit.*, h. 38.

1) Metode Uji Beda Taraf Sukar (*Test of b Difference*)

Camilli & Shepard (1994) mengemukakan bahwa untuk mendeteksi bias butir menggunakan metode uji beda taraf sukar, terlebih dahulu harus ditentukan *standard error* dari \hat{b}_F dan \hat{b}_R . Dengan \hat{b}_F merupakan estimasi parameter taraf sukar kelompok fokal dan \hat{b}_R merupakan estimasi parameter taraf sukar kelompok referensi. Bila S_F dan S_R berturut-turut merupakan *standard error* dari \hat{b}_F dan \hat{b}_R , maka *standard error* perbedaan taraf sukar $\Delta\hat{b} = \hat{b}_F - \hat{b}_R$ dilambangkan dengan: $S_{\Delta\hat{b}} = \sqrt{S_F^2 + S_R^2}$ (2.38)

Statistik yang digunakan untuk menguji signifikansi uji beda taraf sukar butir adalah: $d = \frac{\Delta\hat{b}}{S_{\Delta\hat{b}}}$ (2.39)

Dengan d merupakan aproksimasi sebuah unit variabel pada distribusi normal. Tabel probabilitas distribusi normal digunakan untuk menguji level signifikansi sesuai dengan hipotesis nol yang dirumuskan, $H_0: \Delta\hat{b}=0$.

Parameter b dapat diestimasi menggunakan model L1P, L2P, atau L3P. Mengingat pada model Rasch parameter daya pembeda (a) diasumsikan konstan, selanjutnya Camilli & Shepard (1994) merekomendasikan untuk mendeteksi *DIF* lebih akurat agar menggunakan model L2P atau L3P.

Metode Uji Beda Taraf Sukar ini memiliki keunggulan, yaitu proses perhitungannya sangat sederhana. Sehingga analisis dapat dilakukan secara

manual menggunakan program Excel yang cukup sederhana. Di sisi lain, dibalik kesederhanaan dalam perhitungan tersebut, ada kemungkinan metode Uji Beda Taraf Sukat tersebut kurang cermat dalam perhitungan pendeteksian bias butir. Hal ini disebabkan oleh penggunaan parameter dalam perhitungan hanya melibatkan taraf sukar butir saja. Sehingga metode tersebut kemungkinan tidak dapat mendeteksi butir yang mengandung bias butir tidak seragam (*Non-Uniform*), karena parameter daya pembeda dianggap konstan. Apabila parameter daya pembeda diasumsikan konstan, berarti kecuraman grafik fungsi $P(\theta)$ dianggap sama pula, sehingga bias butir *Non-Uniform* kemungkinan tidak terdeteksi.

Langkah-langkah untuk mendeteksi bias butir menggunakan metode Uji Beda Taraf Sukat adalah sebagai berikut.

- a) Mengestimasi parameter taraf sukar butir kelompok fokal dan kelompok referensi secara terpisah, menggunakan program BILOG-MG.
- b) Menentukan konstanta penyetaraan parameter antara grup fokal dan grup referensi menggunakan program Excel 2010.
- c) Menyetarakan parameter taraf sukar butir kelompok fokal dan referensi menggunakan program Excel 2010.
- d) Menghitung nilai $d = \frac{\Delta \hat{b}}{S_{\Delta \hat{b}}}$, apabila nilai d lebih besar daripada nilai kritis $z=1,96$ atau lebih kecil daripada nilai kritis $z=-1,96$ pada taraf signifikansi $\alpha = 0,05$; maka disimpulkan butir ke- i memuat bias butir.

2) Metode Khi-Kuadrat Lord (*Lord's Chi-Square*)

Untuk mendeteksi bias butir menggunakan metode Khi-Kuadrat Lord digunakan persamaan sebagai berikut.⁷⁹

$$\chi_i^2 = v_i' \Sigma_i^{-1} v_i \dots\dots\dots (2.40)$$

Keterangan:

- i : butir soal ke- i , $i=1, 2, 3, \dots, n$
- v_i : Vektor perbedaan pada pengestimasiian parameter butir ke- i antara kelompok fokal dan kelompok referensi
- v_i' : Transpose matriks vektor perbedaan pada pengestimasiian parameter butir ke- i antara kelompok fokal dan kelompok referensi
- Σ_i^{-1} : Invers jumlah matriks varians-kovarians butir ke- i antara kelompok fokal dan kelompok referensi

Dengan:

$$\Sigma_i^{-1} = \begin{bmatrix} \sigma_a^2 & \sigma_{ab} & \sigma_{ac} \\ \sigma_{ab} & \sigma_b^2 & \sigma_{bc} \\ \sigma_{ac} & \sigma_{bc} & \sigma_c^2 \end{bmatrix}, v_i = \begin{bmatrix} a_F - a_R \\ b_F - b_R \\ c_F - c_R \end{bmatrix}, v_i' = [a_F - a_R \quad b_F - b_R \quad c_F - c_R]$$

⁷⁹ Lord, F. M., *Applications of Item Response Theory to Practical Testing Problem* (New Jersey: Lawrence Erlbaum Associates Publishers, 1980), h. 223.

Keuntungan menggunakan metode χ^2 Lord ini adalah metode ini sensitif untuk mendeteksi bias butir yang tidak konsisten (Non-Uniform). Selain itu, metode ini juga memperbolehkan mengestimasi parameter butir dan kemampuan peserta itu menggunakan program BILOG-MG yang sekarang ini sudah banyak digunakan orang. Dalam penelitian ini deteksi bias dengan metode khi-kuadrat Lord dilakukan dengan menggunakan program Excel 2010.

Seock-Ho Kim & Allan S. Cohen (1995) menyatakan bahwa metode Khi-Kuadrat Lord memiliki akurasi yang tinggi untuk mendeteksi bias butir, sepanjang hasil estimasi terhadap matriks varians kovarian presisi. Sebab salah satu kelemahan metode Khi-Kuadrat Lord, adalah kurang akuratnya estimasi terhadap matriks varians kovarians. Untuk meningkatkan akurasi estimasi terhadap matriks varians kovarian disarankan menggunakan program aplikasi komputer.⁸⁰

Langkah-langkah untuk mendeteksi bias butir menggunakan metode Khi-Kuadrat Lord adalah sebagai berikut.

- a) Mengestimasi parameter butir dan matriks variansi kovariansi kelompok focal dan kelompok referensi secara terpisah menggunakan program BILOG-MG.

⁸⁰ Seock-Ho Kim & Allan S. Cohen, A Comparison of Lord's Chi-Square, Raju's Area Measures, and the Likelihood Ratio Test on Detection of Differential Item Functioning. *Journal Applied of Measurement in Education* 8, 4, 291-312 (diakses 5 Maret 2016), h. 308.

- b) Menentukan konstanta penyetaraan parameter butir antara kelompok fokal dan kelompok referensi dengan metode Rerata dan Sigma menggunakan program Excel 2010.
- c) Menyetarakan skala parameter butir kelompok fokal dan kelompok referensi, dengan metode Rerata dan Sigma menggunakan program Excel 2010.
- d) Menghitung nilai χ^2 menggunakan Program Excel 2010 dan membandingkan harga χ^2 dengan nilai kritis $\chi^2=7,81472$ derajat kebebasan $dk=3$ (karena menggunakan model L3P) pada taraf signifikansi 0,05. Jika χ^2 hitung lebih besar dari harga kritis χ^2 pada taraf signifikansi 0,05, maka disimpulkan butir tersebut mengandung bias butir. Dan sebaliknya, jika χ^2 hitung lebih kecil dari harga kritis χ^2 pada taraf signifikansi 0,05, maka disimpulkan butir tersebut tidak mengandung bias butir.

3) Metode Distribusi Sampling Empiris (*Empirical Sampling Distributions*)

Camilli & Shepard (1994) mengemukakan bahwa metode lain dalam *IRT* yang dapat digunakan untuk mendeteksi bias butir adalah metode distribusi sampling empiris. Dalam metode distribusi sampling empiris, kelompok fokal dan kelompok referensi keduanya secara acak random dipisahkan menjadi dua bagian subsampel F1, F2, R1, dan R2 serta

mendapat perlakuan yang sama seperti kelompok fokal dan kelompok referensi. Karena kedua subsampel ini disusun secara acak, semestinya tidak terdapat muatan bias butir. Pada teori responsi butir unidimensi, bias butir dinyatakan sebagai perbedaan peluang menjawab benar suatu butir soal antara kelompok referensi dan kelompok fokal. Indeks bias butir yang dihasilkan terjadi karena kedua kelompok menunjukkan kesalahan sampling. Nilai ekstrem dari analisis F1-F2 dan R1-R2 ditetapkan sebagai nilai kritis untuk menandai adanya muatan bias butir.⁸¹

Secara matematis, perbedaan peluang menjawab benar suatu tes yang diberikan kepada peserta yang kemampuannya sama, namun berasal dari kelompok yang berbeda dirumuskan sebagai berikut.

$$\Delta P_i = P_R(\theta_j) - P_F(\theta_j) \dots\dots\dots (2.41)$$

Keterangan:

- i : butir soal ke-i, i=1, 2, 3,....., n
- ΔP_i : Perbedaan peluang menjawab benar butir ke-i, peserta ke-j kelompok fokal dan kelompok referensi dengan kemampuan θ .
- $P_F(\theta_j)$: Peluang menjawab benar butir ke-i, peserta ke-j kelompok fokal dengan kemampuan θ .
- $P_R(\theta_j)$: Peluang menjawab benar butir ke-i, peserta ke-j kelompok referensi dengan kemampuan θ .

⁸¹ Camilli, G., & Shepard, L. A., *op. cit.* hh. 73-74.

Lebih lanjut Camilli & Shepard (1994) menyatakan bahwa untuk menghitung indeks perbedaan peluang menjawab benar suatu butir soal dari dua kelompok referensi dan kelompok fokal yang kurva karakteristiknya *uniform* (konsisten/tidak berpotongan) digunakan rumus *Signed Probability Difference (SPD)* sebagai berikut.

$$SPD - \theta = \frac{\sum_{j=1}^{n_p} \Delta P(\theta_j)}{n_p} \dots\dots\dots(2.41.a)$$

Dengan n_p merupakan banyaknya peserta tes pada kelompok p.

Sedangkan untuk menghitung indeks perbedaan peluang menjawab benar suatu butir soal dari dua kelompok referensi dan kelompok fokal yang kurva karakteristiknya *non-uniform* (tidak konsisten/berpotongan) digunakan rumus *Unsigned Probability Difference (UPD)* sebagai berikut.

$$UPD - \theta = \sqrt{\frac{\sum_{j=1}^{n_p} (\Delta P(\theta_j))^2}{n_p}} \dots\dots\dots(2.41.b)$$

Langkah-langkah untuk mendeteksi bias butir menggunakan metode Distribusi Sampling Empiris adalah sebagai berikut.

- a) Mengestimasi parameter butir dan kemampuan kelompok fokal dan kelompok referensi secara terpisah menggunakan program BILOG-MG.
- b) Menentukan konstanta penyetaraan parameter antara grup fokal dan grup referensi, dengan metode Rerata dan Sigma menggunakan program Excel 2010.

- c) Menyetarakan skala parameter butir kelompok fokal dan kelompok referensi.
- d) Menghitung indeks bias butir kelompok referensi dan fokal (δ) menggunakan rumus $SPD - \theta = \frac{\sum_{j=1}^{n_p} \Delta P(\theta_j)}{n_p}$ untuk bias butir jenis *uniform* (konsisten) dan $UPD - \theta = \sqrt{\frac{\sum_{j=1}^{n_p} (\Delta P(\theta_j))^2}{n_p}}$ untuk bias butir jenis *non-uniform* (tidak konsisten). Dengan $\Delta P_i = P_R(\theta_j) - P_F(\theta_j)$.
- e) Memisahkan secara acak masing-masing kelompok referensi dan kelompok fokal menjadi subkelompok R1 dan R2, serta F1 dan F2.
- f) Mengestimasi parameter butir subkelompok referensi (R1 dan R2) dan fokal (F1 dan F2) secara terpisah.
- g) Menentukan konstanta penyetaraan parameter masing-masing subkelompok referensi (R1 dan R2) dan subkelompok fokal (F1 dan F2).
- h) Menghitung nilai indeks bias butir subkelompok R1 dan R2 (δ_1), serta nilai indeks bias butir subkelompok F1 dan F2 (δ_2).
- i) Menentukan nilai maksimum dari ($\delta_1; \delta_2$).
- j) Membandingkan nilai δ dengan nilai maksimum dari ($\delta_1; \delta_2$). Jika nilai δ lebih besar daripada nilai maksimum dari ($\delta_1; \delta_2$), maka butir ke-i dinyatakan memuat bias butir. Sebaliknya bila nilai δ lebih kecil daripada nilai maksimum dari ($\delta_1; \delta_2$), maka butir ke-i dinyatakan tidak memuat bias butir.

e. Perbandingan Sensitivitas Metode DIF

Membandingkan sensitivitas masing-masing metode *DIF* dilakukan melalui dua tahap analisis sebagai berikut: (1) tahap pertama, menguji perbedaan rerata jumlah butir yang mengandung bias butir pada ketiga metode *DIF* menggunakan anava satu jalur; (2) tahap kedua, apabila secara statistika terbukti terdapat perbedaan rerata jumlah butir yang mengandung bias butir pada ketiga metode *DIF*, maka dilanjutkan dengan uji *Post Hoc*, menggunakan uji Tukey (karena jumlah sampel yang diuji sama).

B. HASIL PENELITIAN YANG RELEVAN

Penelitian tentang perbandingan sensitivitas metode *DIF* telah dilakukan dalam beberapa riset sebelumnya. Kajian terhadap hasil penelitian yang relevan merupakan kajian empirik terkait dengan metode pendeteksian bias butir yang menunjukkan adanya indikasi perbedaan sensitivitas metode *DIF* terhadap pendeteksian bias butir.

Berdasarkan studi literatur yang telah dilakukan terhadap hasil-hasil penelitian yang terkait dengan metode pendeteksian bias butir, baik dalam bentuk jurnal maupun disertasi umumnya mengkaji permasalahan yang berbeda-beda yang dilakukan di tempat yang berbeda pula. Berikut ini disajikan ringkasan hasil penelitian yang dianalisis dalam bentuk disertasi.

Tabel 2.1 Ringkasan Studi terdahulu dan studi yang akan dilakukan

Peneliti	Jenis Data	Kajian
Heri Retnawati (2003)	Dikotomus (Skor Seleksi Masuk SLTP)	Membandingkan 3 metode <i>DIF</i> : 1) metode Khi- Kuadrat dari Lord, 2) metode perbandingan kemungkinan, dan 3) metode Mantel-Haenszel.
Badrun Kartowagiran (2005)	Dikotomus (Skor UN Matematika SMP 2003)	Membandingkan 4 metode <i>DIF</i> yaitu: (1) <i>ICC</i> , (2) Luasan menurut Raju, (3) Khi Kuadrat Lord, dan (4) Tes Rasio Kebolehjadian. Dalam penelitian itu, masing-masing metode <i>DIF</i> tidak dibandingkan sensitivitasnya berdasarkan metode penyetaraan yang digunakan.
Wardani Rahayu (2007)	Dikotomus (Skor UN Matematika SMP 2004)	Membandingkan metode <i>linking</i> : (1) rerata dan sigma, (2) tegar rerata dan sigma, dan (3) kurva karakteristik, pada metode <i>DIF Lord Chi Square</i> , untuk mendeteksi butir <i>false positive</i> .
Effendi (2011)	Dikotomus (Skor UN Kimia SMA 2008)	Membandingkan metode <i>DIF</i> : (1) Luasan Raju (2) <i>Lord Chi Square</i> , dan (3) Tes Rasio Kebolehjadian, pada metode <i>linking</i> kurva karakteristik (<i>ICC</i>), untuk mendeteksi butir <i>crossing DIF</i> .

Peneliti	Jenis Data	Kajian
Idrus Alwi (2011)	Dikotomus (Skor tes matematika kls IX smt. 1 MTS)	Membandingkan kepekaan model Mantel-Haenszel dan model Rasch pada sampel ukuran responden 400 dan 200.
I Wayan Widana (Rancangan disertasi 2015)	Dikotomus (Skor UN Matematika SMA IPA 2015)	Membandingkan sensitivitas metode <i>DIF</i> Uji Beda Taraf Sukar, Khi-Kuadrat Lord, dan Distribusi Sampling Empiris (berdasarkan IRT model L3P) untuk mendeteksi bias butir.

Sesuai dengan ringkasan hasil penelitian pada tabel 2.1 dapat dipaparkan hasil kajian terhadap hasil penelitian terdahulu sebagai berikut.

Heri Retnawati (2003) dalam penelitiannya yang berjudul “Keberfungsian Butir Diferensial pada Perangkat Tes Seleksi Masuk SLTP Mata Pelajaran Matematika”, mengkaji 3 metode *DIF* yaitu: 1) metode Khi-Kuadrat dari Lord, 2) metode perbandingan kemungkinan, dan 3) metode Mantel-Haenszel. Hasil penelitian menyimpulkan bahwa metode Mantel-Haenszel merupakan metode yang paling banyak menghasilkan butir soal memuat bias butir, namun belum diketahui besarnya kesalahan dalam pendeteksian bias butir.⁸²

⁸² Heri Retnawati, Keberfungsian Butir Diferensial pada Perangkat Tes Seleksi Masuk SLTP Mata Pelajaran Matematika. Tesis (Yogyakarta: UNY), h. i.

Badrun Kartowagiran (2005) dalam penelitiannya yang berjudul “Perbandingan Berbagai Metode untuk Mendeteksi Bias Butir”, membandingkan 4 metode *DIF*. Hasil penelitian ini menunjukkan bahwa di antara 4 metode *DIF* yang dibandingkan yaitu: (1) *ICC*, (2) Luasan menurut Raju, (3) Khi Kuadrat Lord, dan (4) Tes Rasio Kebolehjian, dinyatakan bahwa metode deteksi *DIF* yang paling sensitif untuk tes UAN Matematika SMP tahun 2003 adalah metode tes rasio kebolehjian, baik berdasarkan jumlah butir yang memuat bias butir maupun berdasarkan hasil analisis faktor konfirmatori.⁸³

Wardani Rahayu (2007) dalam penelitiannya yang berjudul “Pengaruh Metode *Linking* terhadap Banyak Butir *False Positive* pada Pendeteksian *DIF* Berdasarkan Teori Responsi Butir”, mengkaji keakuratan pendeteksian *DIF* menggunakan metode *Lord's Chi-Square* dengan membandingkan metode *linking*: (1) rerata dan sigma, (2) tegar rerata dan sigma, dan (3) kurva karakteristik, pada metode *DIF Lord Chi Square*, untuk mendeteksi butir *false positive* (butir yang diduga tidak *DIF*, ternyata *DIF*). Data yang digunakan adalah data dikotomus berupa hasil UN SMP mata pelajaran matematika tahun 2004. Hasil penelitian menyatakan bahwa pada pendeteksian *DIF* menggunakan metode *Lord's Chi-Square*, bahwa metode *linking* kurva karakteristik dan metode rerata dan sigma lebih akurat daripada metode

⁸³ Badrun Kartowagiran, *op. cit.*, h. v.

tegar rerata dan sigma. Metode kurva karakteristik memiliki keakuratan yang sama dengan metode rerata dan sigma bila menggunakan metode *DIF Lord's Chi-Square*. Lebih lanjut Wardani menyarankan agar menggunakan metode kurva karakteristik atau metode rerata dan sigma untuk penyamaan skala berdasarkan teori responsi butir.⁸⁴

Effendi (2011) dalam penelitiannya yang berjudul "Pendeteksian *Crossing DIF*: Suatu Komparasi Metode Luasan Menurut Raju Chi-Square Menurut Lord, dan Tes Rasio Kebolehjadian", menyatakan bahwa (1) metode luasan menurut Raju lebih sensitif dalam pendeteksian *Crossing DIF* daripada metode tes rasio kebolehjadian, (2) metode *chi-square* menurut Lord lebih sensitif dalam pendeteksian *Crossing DIF* daripada metode tes rasio kebolehjadian, dan (3) metode luasan menurut Raju lebih sensitif dalam pendeteksian *Crossing DIF* daripada metode *chi-square* menurut Lord.⁸⁵ Kesimpulan itu adalah wajar karena masing-masing metode pendeteksi bias butir memiliki teknik yang berbeda-beda.

Idrus Alwi (2011) dalam penelitiannya yang berjudul "Perbandingan Kepekaan Model Mantel Haenszel dan Model Rasch dalam Mendeteksi Keberbedaan Fungsi Butir Ditinjau dari Ukuran Responden". Berdasarkan

⁸⁴ Wardani Rahayu, Pengaruh Metode *Linking* terhadap Banyak Butir False Positive pada Pendeteksian *DIF* Berdasarkan Teori Responsi Butir. *Disertasi* (Jakarta: UNJ, 2007), hh. xxi-xxii.

⁸⁵ Effendi, Pendeteksian *Crossing DIF*: Suatu Komparasi Metode Luasan Menurut Raju Chi-Square Menurut Lord, dan Tes Rasio Kebolehjadian. *Disertasi*. (Jakarta: UNJ, 2011), hh. 6-7.

hasil pengujian terhadap metode deteksi keberbedaan fungsi butir, didapatkan kesimpulan: hasil perbandingan antara metode deteksi keberbedaan fungsi butir model Rasch dengan model Mantel-Haenszel menyatakan bahwa tingkat kepekaan deteksi keberbedaan fungsi butir model Rasch lebih tinggi dibandingkan dengan metode deteksi keberbedaan fungsi butir model Mantel-Haenszel baik menggunakan ukuran responden 400 maupun 200.⁸⁶

Berdasarkan kajian empiris di atas, tampak bahwa masing-masing penelitian memiliki sasaran kajian yang berbeda antara yang satu dengan lainnya. Demikian pula kajian yang dilakukan pada penelitian ini, menguji sensitivitas metode *DIF* yang berbeda dengan studi yang telah dilakukan sebelumnya. Adapun metode *DIF* yang diteliti dan diperbandingkan sensitivitasnya dalam penelitian ini adalah: (1) metode Distribusi Sampling Empiris, (2) metode Khi-Kuadrat Lord, dan (3) metode Uji Beda Taraf Sukar Butir. Ketiga metode *DIF* tersebut berdasarkan teori responsi butir. Dengan demikian, hasil penelitian ini diharapkan dapat menambah referensi kajian tentang sensitivitas metode *DIF*.

⁸⁶ Alwi Idrus, Perbandingan Kepekaan Model Mantel Haenszel dan Model Rasch dalam Mendeteksi Keberbedaan Fungsi Butir Ditinjau Dari Ukuran Responden. *Disertasi*. (Jakarta: UNJ, 2011), hh. Viii.

C. KERANGKA TEORETIK

1. Perbandingan sensitivitas untuk mendeteksi Bias Butir antara metode Distribusi Sampling Empiris (DSE) dan Khi-Kuadrat Lord

Pada metode *DIF* Distribusi Sampling Empiris, untuk mendeteksi bias butir menggunakan selisih probabilitas menjawab benar dari sebuah butir soal yang diberikan kepada peserta dengan kemampuan yang sama pada kelompok yang berbeda. Secara teoretis, keunggulan metode DSE dibandingkan dengan metode Khi-Kuadrat Lord, sebagaimana diungkapkan oleh Shepard, dkk (1994) bahwa untuk mendeteksi bias butir dilakukan melalui replikasi analisis, menggunakan sampel berpasangan yang ekuivalen random.

Replikasi analisis dilakukan dengan terlebih dahulu membagi dua masing-masing kelompok referensi dan fokal secara random, sebagai subsampel R1–R2, dan F1-F2. Shepard, dkk memilih secara random R1 dan R2 serta F1 dan F2, kemudian memberi perlakuan yang sama seperti kelompok referensi dan kelompok fokal sebelumnya. Karena kelompok ini dibentuk berdasarkan pemilihan secara random, seharusnya tidak ada bias pada butir. Namun, indeks bias butir masih diperoleh untuk kedua kelompok yang ditunjukkan oleh varians akibat *sampling error*. Dengan membandingkan indeks hasil analisis R1, F1 dengan hasil analisis R2, F2, Shepard, dkk dapat mengevaluasi stabilitas indeks antarsampel. Nilai

maksimum dari analisis R1-R2 dan F1-F2 dipilih sebagai nilai kritis untuk signifikansi *DIF*.⁸⁷

Hasil analisis bias butir menggunakan metode DSE dipengaruhi oleh 3 komponen yaitu (1) perbedaan probabilitas menjawab benar kelompok referensi dan kelompok fokal (δ), (2) perbedaan probabilitas menjawab benar subkelompok referensi (δ_1), dan (3) perbedaan probabilitas menjawab benar subkelompok fokal (δ_2). Pertama, dihitung indeks bias butir kelompok

referensi dan fokal (δ) menggunakan rumus $SPD - \theta = \frac{\sum_{j=1}^{n_p} \Delta P(\theta_j)}{n_p}$ untuk bias

butir jenis *uniform* (konsisten) dan $UPD - \theta = \sqrt{\frac{\sum_{j=1}^{n_p} (\Delta P(\theta_j))^2}{n_p}}$ untuk bias butir

jenis *non-uniform* (tidak konsisten). Dengan $\Delta P_i = P_R(\theta_j) - P_F(\theta_j)$.

Selanjutnya pada masing-masing subkelompok dihitung $\Delta P_F = P_{F1}(\theta_j) -$

$P_{F2}(\theta_j) = \delta_1$ dan $\Delta P_R = P_{R1}(\theta_j) - P_{R2}(\theta_j) = \delta_2$, serta di tentukan nilai

maksimum dari ($\delta_1; \delta_2$). Bila nilai δ lebih besar daripada nilai maksimum dari

($\delta_1; \delta_2$), maka butir ke-i dinyatakan sebagai butir bias. Sebaliknya jika nilai δ

lebih kecil daripada nilai maksimum dari ($\delta_1; \delta_2$), maka butir ke-i dinyatakan

sebagai butir tidak bias.

⁸⁷ Camilli G., and Shepard L., *op.cit.*, hh. 73-74.

Oleh karena masing-masing subsampel kelompok referensi (R1-R2) dan fokal (F1-F2) berasal dari kelompok referensi dan fokal yang sama serta dipilih secara acak, maka semestinya δ_1 dan δ_2 nilainya cukup kecil. Kecilnya nilai δ_1 dan δ_2 diakibatkan oleh kecilnya perbedaan nilai estimasi parameter butir dan nilai θ yang dibangkitkan dari subkelompok sampel, sehingga perbedaan $P(\theta)$ pada subkelompok sampel cukup kecil. Dengan kata lain, dalam subkelompok referensi (R1-R2) dan subkelompok fokal (F1-F2) masing-masing tidak terjadi bias butir. Berbeda dengan perhitungan δ pada kelompok referensi dan kelompok fokal yang melibatkan jumlah sampel yang lebih besar (dua kali lipat jumlah subkelompok sampel) sehingga memungkinkan perbedaan nilai estimasi parameter butir yang lebih besar, akibatnya perbedaan nilai $P(\theta)$ pada kelompok referensi dan fokal cukup besar, sehingga nilai δ lebih besar daripada nilai maksimum δ_1 dan δ_2 . Dalam kondisi demikian dinyatakan sebuah butir ke- i memuat bias butir.

Pada metode Khi-Kuadrat Lord, signifikansi terhadap bias butir diperhitungkan melalui perbedaan estimasi parameter butir dan matriks varians kovarians kelompok referensi dan fokal, tanpa adanya replikasi analisis seperti dalam metode DSE. Nilai χ^2 dipengaruhi oleh perbedaan estimasi parameter butir dan besarnya elemen-elemen matriks varians kovarian kelompok referensi dan fokal. Varians merupakan ukuran variabilitas distribusi suatu variabel yang dilambangkan dengan σ_x^2 . Apabila μ merupakan

rerata dari distribusi suatu variabel, maka $x - \mu$ merupakan penyimpangan suatu data terhadap rataannya. Oleh karena varians merupakan rerata kuadrat penyimpangan, maka nilai σ_x^2 akan lebih kecil untuk kelompok data yang dekat dengan μ dibandingkan dengan kelompok data yang jauh dari μ . Dengan kata lain, jika nilai-nilai x cenderung terkonsentrasi di dekat rataannya, maka variansinya kecil mendekati nol. Sedangkan jika x jauh dari rataannya maka variansinya besar. Selanjutnya kovarians menyatakan kekuatan hubungan dua variabel. Selain mengukur besarnya kekuatan hubungan di antara dua variabel, kovarians juga menentukan arah hubungan dari kedua variabel tersebut. Nilai kovarians dibedakan menjadi 3, yaitu: (a) bertanda positif menyatakan bahwa semakin tinggi nilai variabel x , semakin tinggi pula nilai variabel y (menyatakan hubungan searah); (b) bertanda negatif menyatakan bahwa semakin tinggi nilai variabel x , semakin rendah nilai variabel y (menyatakan hubungan berlawanan arah); dan (c) mendekati nol berarti bahwa kedua variabel memiliki hubungan yang sangat lemah (tidak saling mempengaruhi).

Untuk menganalisis bias butir menggunakan metode Khi-Kuadrat Lord, secara matematis dinyatakan dalam bentuk perkalian matriks: $\chi_i^2 = v_i' \Sigma_i^{-1} v_i$. Sebuah butir dinyatakan memuat bias butir bila nilai $\chi_{hitung}^2 \geq \chi_{tabel}^2 (dk=3; \alpha=0.05)$. Semakin besar elemen-elemen pada matriks varians

kovarians maka elemen-elemen inversnya semakin kecil, dan sebaliknya.

Misalkan $\Sigma A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, dan $\Sigma B = k \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$. Dengan

$\Sigma B = \Sigma kA$, maka dapat dibuktikan bahwa $\Sigma^{-1}B = \frac{1}{k}\Sigma^{-1}A$. Apabila nilai

perbedaan masing-masing parameter butir ke- i diasumsikan konstan, dapat

disimpulkan bahwa semakin besar nilai elemen-elemen penjumlahan matriks

varians-kovarians kelompok referensi dan focal maka semakin kecil nilai χ^2 .

Selanjutnya, apabila invers matriks varians kovarians diasumsikan konstan,

maka semakin besar perbedaan estimasi parameter butir pada kelompok

referensi dan focal maka semakin besar pula nilai χ^2 . Apabila perbedaan

masing-masing parameter butir ke- i k kali perbedaan masing-masing

parameter butir ke- j , maka nilai $\chi_j^2 = k^2 \cdot \chi_i^2$. Dengan demikian, nilai χ^2

semakin besar bila perbedaan masing-masing parameter butir semakin besar

dan elemen-elemen jumlah matriks varians kovarians semakin kecil.

Dari uraian di atas dapat dilihat bahwa di antara kedua metode *DIF*

pendeteksi bias butir tersebut, terdapat perbedaan teknik dalam

perhitungannya. Metode DSE dalam pendeteksian bias butir menggunakan

selisih probabilitas menjawab benar dari sebuah butir soal yang diberikan

kepada peserta dengan kemampuan yang sama pada kelompok berbeda.

Artinya, dalam perhitungan tersebut juga mempertimbangkan parameter θ

(kemampuan peserta). Seharusnya, apabila suatu butir soal tidak bias, maka

$P(\theta)$ untuk kedua kelompok sama, atau kurva karakteristik butir itu berimpit yang mengakibatkan $\Delta P_i = 0$. Perhitungan tersebut lebih cermat bila dibandingkan dengan metode Khi-Kuadrat Lord, karena dalam perhitungannya hanya menggunakan selisih estimasi parameter butir yang dinyatakan dalam bentuk perkalian matriks. Dalam perhitungan metode Khi-Kuadrat Lord tidak mempertimbangkan parameter θ (kemampuan peserta), dan tidak melakukan replikasi analisis. Padahal sesuai dengan definisi bias butir yaitu butir yang memberikan peluang tidak sama untuk menjawab benar terhadap butir tertentu pada kemampuan yang sama hanya karena berasal dari kelompok yang berbeda. Sehingga ada kemungkinan metode Khi-Kuadrat Lord kurang cermat mendeteksi bias butir karena mengabaikan parameter θ , dan tidak melakukan replikasi analisis.

Memperhatikan pendapat Camilli dan Shepard (1994) yang menyatakan bahwa pendeteksian bias butir akan lebih cermat bila mempertimbangkan parameter kemampuan dan parameter butir secara lengkap. Dengan demikian metode *DIF* Distribusi Sampling empiris diduga lebih sensitif untuk mendeteksi bias butir, bila dibandingkan dengan metode Khi-Kuadrat Lord.

2. Perbandingan sensitivitas untuk mendeteksi Bias Butir antara metode Khi-Kuadrat Lord dan Uji Beda Taraf Sukar

Pada metode UBTS, perhitungan dalam analisis bias butir sangat sederhana hanya didasarkan pada perbedaan estimasi parameter taraf sukar ($\Delta\hat{b}$) dan *standard error* taraf sukar pada masing-masing kelompok referensi (S_R) dan kelompok fokal (S_F). Sedangkan parameter daya pembeda (a) diasumsikan konstan (sama) untuk semua butir. Demikian pula parameter kebetulan menjawab betul (c) diasumsikan nol. Sehingga kedua parameter daya pembeda dan kebetulan menjawab betul tidak dipertimbangkan dalam perhitungan *DIF*. Secara matematis, perhitungan indeks bias butir pada metode UBTS dinyatakan dalam bentuk: $d = \frac{\Delta\hat{b}}{\sqrt{S_R^2 + S_F^2}}$. Nilai d akan bertambah besar apabila nilai $\Delta\hat{b}$ semakin besar dan nilai $\sqrt{S_R^2 + S_F^2}$ semakin kecil. Sebaliknya nilai d semakin kecil bila $\Delta\hat{b}$ semakin kecil dan nilai $\sqrt{S_R^2 + S_F^2}$ semakin besar. Apabila nilai $d > 1,96$ atau $d < -1,96$ maka butir ke- i dinyatakan sebagai butir yang bias.

Dalam metode Khi-Kuadrat Lord, untuk mendeteksi bias butir melibatkan semua parameter butir. Perbedaan nilai estimasi parameter a (daya pembeda) mengakibatkan perbedaan kecuraman kurva karakteristik butir. Semakin besar nilai a , maka kurva karakteristik butir semakin curam, dan semakin kecil nilai a maka kurva karakteristik butir semakin landai.

Perbedaan nilai estimasi parameter b , mengakibatkan terjadi pergeseran kurva karakteristik butir ke kiri atau ke kanan. Semakin besar nilai b (butir tersebut semakin sulit), maka posisi kurva karakteristik butir semakin ke kanan. Sedangkan perbedaan nilai estimasi parameter c , mengakibatkan pergeseran kurva karakteristik butir ke atas atau ke bawah. Semakin besar nilai c (tebakan semakin tinggi), maka posisi kurva karakteristik butir semakin ke atas.

Berbeda dengan metode UBTS yang mengasumsikan parameter daya pembeda sama dan faktor tebakan diasumsikan nol. Asumsi-asumsi tersebut mengakibatkan perbedaan posisi kurva karakteristik kurva pada kelompok referensi dan focal tidak terdeteksi dengan baik, akibat perbedaan estimasi a dan c tidak terdeteksi. Dengan demikian metode UBTS kurang peka untuk mendeteksi bias butir bila dibandingkan dengan metode Khi-Kuadrat Lord. Semakin lengkap parameter butir dan parameter kemampuan yang digunakan dalam pendeteksian bias butir semakin akurat (sensitif) pula metode *DIF* tersebut digunakan untuk mendeteksi bias butir. Hal itu dapat dipahami karena semakin banyak parameter yang digunakan dalam menghitung $P(\theta)$, maka kurva karakteristik butir yang diperoleh semakin presisi, ditinjau dari kecuraman (*slope*), taraf sukar (*difficulty*), dan kebetulan menjawab betul (*pseudoguessing*). Dengan demikian diduga metode Khi-

Kuadrat Lord lebih sensitif untuk mendeteksi bias butir bila dibandingkan dengan metode Uji Beda Taraf Sukar.

D. HIPOTESIS PENELITIAN

1. Metode Distribusi Sampling Empiris lebih sensitif daripada metode Khi-Kuadrat Lord untuk mendeteksi bias butir.
2. Metode Khi-Kuadrat Lord lebih sensitif daripada metode Uji Beda Taraf Sukar untuk mendeteksi bias butir.

BAB III

METODOLOGI PENELITIAN

A. Tujuan Penelitian

Secara umum penelitian ini bertujuan untuk mengetahui perbandingan sensitivitas metode *DIF* untuk mendeteksi bias butir. Untuk mendeteksi bias butir yang disebabkan oleh perbedaan gender, dilakukan dengan mengelompokkan siswa menjadi dua kelompok yang berbeda, yaitu peserta didik laki-laki sebagai kelompok referensi dan peserta didik perempuan sebagai kelompok fokal.

Selanjutnya, secara khusus penelitian ini dilakukan untuk mengetahui:

- 1) perbandingan sensitivitas metode *DIF* untuk mendeteksi bias butir pada metode Distribusi Sampling Empiris dan Khi-Kuadrat Lord;
- 2) perbandingan sensitivitas metode *DIF* untuk mendeteksi bias butir pada metode Khi-Kuadrat Lord dan Uji Beda Taraf Sukar.

B. Tempat dan Waktu Penelitian

Penelitian ini dilaksanakan di Provinsi DKI Jakarta. Pemilihan Provinsi DKI Jakarta sebagai tempat penelitian didasarkan atas penilaian BSNP bahwa hasil UN SMA di Provinsi DKI Jakarta termasuk salah satu provinsi yang dinyatakan jujur dalam pelaksanaan UN Tahun Pelajaran 2014/2015. Hal itu dapat dilihat dengan angka indeks integritas UN (IIUN) yang tinggi

pada sekolah-sekolah SMA di Provinsi DKI Jakarta. Hasil tersebut dianggap dapat menggambarkan hubungan antara pola jawaban dengan kemampuan peserta tes mendekati kondisi sesungguhnya.

Penelitian dilakukan mulai bulan November 2015 sampai dengan Maret 2016. Penelitian dilakukan dengan tahapan-tahapan sebagai berikut: 1) penelitian diawali dengan menyusun proposal penelitian dan seminar proposal yang dilaksanakan tanggal 26 November 2015; 2) mempersiapkan *software* pendukung untuk keperluan analisis data dilakukan pada bulan Desember 2015 yaitu program BILOG-MG, SPSS 22.0, dan program Excel 2010; 3) meminta data repon hasil UN Provinsi DKI Jakarta untuk mata pelajaran Matematika Program IPA tahun pelajaran 2014/2015 kepada Pusat Penilaian Pendidikan (Puspendik) Balitbang Kemdikbud sekitar bulan Desember 2015; 4) melakukan analisis data hasil UN Matematika Program IPA dan penyusunan laporan hasil penelitian pada bulan Januari-Maret 2016; 5) melakukan seminar hasil penelitian yang dilakukan tanggal 8 Juni 2016.

C. Metode Penelitian

Penelitian ini termasuk dalam penelitian deskriptif eksploratif, yaitu penelitian yang akan mendeskripsikan kinerja tiga metode *DIF*. Penelitian ini juga mendeskripsikan metode *DIF* yang paling sensitif untuk mendeteksi bias butir. Metode *DIF* yang diperbandingkan dalam penelitian ini adalah metode Uji Beda Taraf Sukar (*Test of Difficulty Difference Method*), metode Khi-

Kuadrat Lord (*Lord's Chi-Square Method*), dan metode Distribusi Sampling Empiris (*Empirical Sampling Distributions Method*).

Selanjutnya, rancangan penelitian yang akan dilakukan dapat digambarkan sebagai berikut.

Tabel 3.1 Rancangan Penelitian

Replikasi ke:	Jumlah Bias Butir		
	Distribusi Sampling Empiris	Khi-Kuadrat Lord	Uji Beda Taraf Sukar
1.			
2.			
3.			
....			
30.			

D. Populasi dan Sampel

Dalam penelitian ini populasi dibagi menjadi dua jenis yaitu populasi peserta tes dan populasi butir. Populasi peserta tes adalah peserta didik SMA Negeri dan Swasta di Provinsi DKI Jakarta peserta UN tahun pelajaran 2014/2015, yang dipilah menjadi populasi peserta didik laki-laki dan peserta didik perempuan. Populasi peserta tes berjumlah 4.869 orang peserta didik, yang terdiri dari 2.114 orang peserta didik laki-laki dan 2.755 peserta didik perempuan. Sampel peserta tes diambil secara acak dari masing-masing populasi yaitu kelompok perempuan sebagai kelompok fokal dan kelompok

laki-laki sebagai kelompok referensi. Jumlah sampel responden yang akan digunakan dalam penelitian ini adalah 2.000 orang. Masing-masing 1.000 orang untuk kelompok referensi dan 1.000 orang untuk kelompok fokal. Hal ini sesuai dengan pendapat Dali S. Naga (1992) yang menyatakan bahwa ukuran sampel minimal pada teori responsi butir sangat tergantung pada model yang digunakan. Untuk model L3P ukuran sampelnya lebih besar daripada model L2P, ada ukuran sampel yang kurang dari 500, ada yang kurang dari 1.000, ada pula yang kurang dari 2.000. Semakin besar ukuran sampel, semakin baik hasil estimasi model respons karena berkaitan dengan uji kecocokan model.⁸⁸

Sedangkan populasi butir dalam penelitian ini adalah butir-butir soal Matematika Program IPA, yang terdapat dalam perangkat tes UN Provinsi DKI Jakarta tahun pelajaran 2014/2015 dalam bentuk paket-paket soal. Sampel butir adalah butir-butir soal UN mata pelajaran Matematika Program IPA sebanyak 40 butir, pada paket dengan kode 1101.

E. Teknik Pengumpulan Data

Data dalam penelitian ini adalah data hasil UN mata pelajaran Matematika Program IPA Provinsi DKI Jakarta tahun pelajaran 2014/2015. Data tersebut diperoleh dari Puspendik Balitbang Kemdikbud. Data tersebut berupa respons peserta didik dalam format *Excel*. Mengingat kebijakan yang

⁸⁸ Dali S. Naga., *op. cit.*, h. 168.

dilakukan oleh Puspendik tidak mengeluarkan kunci jawaban, maka kunci jawaban dibuat oleh peneliti bersama guru-guru Matematika SMA Senior dan dosen Matematika dari Perguruan Tinggi. Respons peserta didik yang awalnya dalam bentuk huruf dikonversi dalam bentuk angka yaitu untuk jawaban benar diberi angka 1 dan jawaban salah diberi angka 0, sehingga dapat dilakukan penskoran.

1) Definisi Konseptual Variabel

Berikut ini akan diuraikan secara singkat definisi konseptual masing-masing variabel.

- a. Sensitivitas metode *DIF* adalah kemampuan suatu metode *DIF* untuk mendeteksi adanya bias butir dalam suatu perangkat tes. Sensitivitas metode *DIF* dapat dilihat dari banyaknya jumlah butir yang terdeteksi mengandung bias butir oleh metode *DIF*.
- b. Metode Uji Beda Taraf Sukar (*Test of Difficulty Difference Method*) merupakan salah satu metode *DIF* pendeteksi bias butir berdasarkan teori responsi butir yang dalam perhitungan analisis bias butir hanya memperhitungkan taraf sukar butir (b). Parameter daya pembeda diasumsikan sama (konstan) dan parameter c dianggap nol. Untuk mendeteksi bias butir digunakan rumus:
$$d = \frac{\Delta \hat{b}}{S_{\Delta \hat{b}}}$$

c. Metode Khi-Kuadrat Lord (*Lord's Chi-Square Method*) adalah salah satu metode *DIF* untuk mendeteksi bias butir berdasarkan teori responsi butir, yang dalam perhitungan untuk mendeteksi bias butir menggunakan parameter butir secara lengkap. Untuk mendeteksi bias butir, diperhitungkan menggunakan rumus: $\chi_i^2 = v_i' \Sigma_i^{-1} v_i$.

d. Metode Distribusi Sampling Empiris (*Empirical Sampling Distributions Method*) juga merupakan salah satu metode *DIF* untuk mendeteksi bias butir berdasarkan teori responsi butir yang dalam perhitungan untuk mendeteksi bias butir menggunakan selisih peluang menjawab benar antara peserta kelompok fokal dan kelompok referensi untuk masing-masing butir. Untuk mendeteksi bias butir dilakukan melalui dua tahap. Tahap pertama, menghitung indek bias butir antara kelompok fokal dan referensi (R dan F) menggunakan rumus: $SPD - \theta = \frac{\sum_{j=1}^{n_p} \Delta P(\theta_j)}{n_p}$ untuk bias

butir *uniform* (konsisten) dan $UPD - \theta = \sqrt{\frac{\sum_{j=1}^{n_p} (\Delta P(\theta_j))^2}{n_p}}$ untuk bias butir

non-uniform (tidak konsisten), dengan: $\Delta P_i = P_R(\theta_j) - P_F(\theta_j)$. Indek bias butir antara kelompok fokal dan referensi dilambangkan dengan δ . Tahap kedua menghitung indeks bias butir subkelompok masing-masing R1 dan R2 dilambangkan dengan δ_1 , serta F1 dan F2 dilambangkan dengan δ_2 . Rumus yang digunakan menghitung δ_1 dan δ_2 sama dengan menghitung

indeks bias butir kelompok fokal dan referensi (R-F). Untuk mendeteksi bias butir, dilakukan dengan cara membandingkan indeks bias butir kelompok referensi dan fokal (δ) dengan nilai maksimum (δ_1, δ_2).

2) Prosedur Penelitian

Prosedur penelitian merupakan langkah-langkah pokok yang dilakukan dalam penelitian. Prosedur penelitian berisikan langkah-langkah umum yang akan dilakukan dalam penelitian. Berikut dipaparkan tahapan-tahapan yang dilakukan dalam penelitian ini.

- a. Mengajukan permohonan kepada Puspendik, untuk memperoleh data hasil UN Mata Pelajaran Matematika Program IPA Provinsi DKI Jakarta tahun pelajaran 2014/2015 untuk Paket Soal dengan kode 1101.
- b. Memilah populasi data yang diperoleh berdasarkan jenis kelamin laki-laki dan perempuan, serta menetapkan kelompok siswa perempuan sebagai kelompok fokal dan kelompok siswa laki-laki sebagai kelompok referensi.
- c. Menguji kecocokan model dengan data yang dihasilkan dari lapangan. Uji kecocokan model dilakukan dengan Program BILOG-MG. Kecocokan butir dengan model logistik tiga parameter (L3P) dapat dilihat dari harga *p-value* pada kolom *CHISQ* pada *output.PH2* yang dijalankan dengan program BILOG-MG. Butir-butir yang memiliki harga *p-value* lebih dari 0,05 maka butir tersebut dianggap cocok dengan model logistik tiga

parameter (L3P). Hanya butir-butir yang cocok dengan model yang dianalisis, butir yang tidak cocok dihilangkan.

- d. Melakukan replikasi pemilihan kelompok sampel peserta tes sebanyak 30 kali, masing-masing untuk kelompok fokal sebanyak 1.000 orang dan kelompok referensi juga 1.000 orang. Dengan demikian akan terdapat 30 kelompok sampel peserta tes.
- e. Pada masing-masing replikasi, parameter butir dan parameter peserta tes, variansi kovariansi butir, fungsi informasi butir pada kelompok referensi dan kelompok fokal diestimasi secara terpisah untuk butir-butir yang dinyatakan cocok dengan model. Estimasi ini dilakukan dengan menggunakan program BILOG-MG. Langkah-langkah yang dilakukan untuk mengestimasi parameter butir dan parameter peserta tes, variansi kovariansi butir, fungsi informasi tes kelompok referensi dan kelompok fokal secara terpisah dengan program BILOG-MG adalah: (1) mengkonversi data dalam bentuk Excel ke bentuk Notepad; (2) membuat sintaks program BILOG-MG; (3) menjalankan program BILOG-MG.
- f. Menyamakan skala daya beda butir (a), taraf sukar butir (b), tingkat kebetulan menjawab betul pada butir (c), dan parameter kemampuan (θ) dari kelompok fokal ke kelompok referensi.

- g. Mendeteksi bias butir dengan tiga metode *DIF* yaitu metode Uji Beda Taraf Sukar, metode Khi-Kuadrat Lord, dan metode Distribusi Sampling Empiris, menggunakan Program Excel 2010.
- h. Membandingkan sensitivitas masing-masing metode *DIF* dilakukan dengan dua tahap, sebagai berikut: (1) tahap pertama, menguji perbedaan rerata jumlah bias butir pada ketiga metode *DIF* menggunakan anava satu jalur; (2) tahap kedua, apabila secara statistika terbukti terdapat perbedaan rerata jumlah bias butir pada ketiga metode *DIF*, maka dilanjutkan dengan uji *Post Hoc*, menggunakan uji Tukey (karena jumlah sampel yang diuji sama untuk masing-masing kelompok).

F. Teknik Analisis Data

1) Uji Persyaratan Analisis

Mengingat analisis data dilakukan menggunakan statistik parameterik, maka terlebih dahulu dilakukan uji persyaratan analisis yaitu uji normalitas dan homogenitas.

a. Uji Normalitas

Uji normalitas data bertujuan untuk menunjukkan bahwa data sampel berasal dari populasi yang berdistribusi normal. Berdasarkan pendapat Schewhart (1988: 53-61) dan Conover (1980: 444) dalam Agung (2006) menyatakan bahwa statistik rerata atau himpunan semua nilai rerata yang

mungkin diobservasi dengan ukuran sampel tertentu yang juga disebut ruang sampel rerata mempunyai pendekatan distribusi normal, berdasarkan Teorema Limit Sentral. Menurut Conover (1980) ukuran sampel minimal adalah 20.⁸⁹ Dengan demikian, pengujian hipotesis statistik tentang rerata dengan jumlah sampel minimal 20 tidak perlu dilakukan, karena menurut Teorema Limit Sentral syarat normalitas telah terpenuhi.

Agung (2004) mengemukakan bahwa Teorema Limit Sentral dinyatakan sebagai berikut: “Jika $X_1, X_2, \dots, \text{ dan } X_n$ sebuah sampel random dari variabel random X dengan $E(X)=\mu$ dan $\text{Var}(X)=\sigma^2$, maka statistik rerata $T_n=T_n(X_1, X_2, \dots, X_n)= (X_1+X_2+\dots+X_n)/n$ berdistribusi normal dengan $E(T)=\mu$ dan $\text{Var}(T)=\frac{\sigma^2}{n}$ jika n mendekati tak terhingga, tidak tergantung pada distribusi variabel random X ”. Berdasarkan teorema tersebut, maka statistik $Z = \frac{T_n - \mu}{\frac{\sigma}{\sqrt{n}}}$ dapat dibuktikan berdistribusi normal standar, dengan $E(Z)=0$ dan $\text{Var}(Z)=1$. Teorema ini juga secara tegas menyatakan bahwa distribusi T tidak tergantung pada distribusi variabel random X . Pernyataan ini mempunyai pengertian bahwa random variabel X tidak perlu dan tidak harus mempunyai distribusi normal untuk memperoleh statistik rerata T yang mempunyai (pendekatan) distribusi normal.⁹⁰

⁸⁹ I Gusti Ngurah Agung, *Statistika Penerapan Model Rerata Sel Multivariat dan Model Ekonometri dengan SPSS* (Jakarta: Yayasan Sad Satria Bhakti, 2006), hh. 83-84.

⁹⁰ I Gusti Ngurah Agung, *Manajemen Penulisan Skripsi, Tesis, dan Disertasi* (Jakarta: PT. Raja Grafindo Persada, 2004), hh. 222-223.

Sejalan dengan pendapat di atas, Wonnacott and Wonnacott dalam Dali S. Naga (2008) menyatakan bahwa persyaratan distribusi probabilitas normal pada pengujian hipotesis statistik tentang rerata tidak peka. Dengan jumlah sampel $n \geq 20$ syarat normalitas tidak lagi berpengaruh pada distribusi pensampelan. Dengan demikian pada pengujian hipotesis statistik untuk sampel yang cukup besar persyaratan normalitas populasi tidak terlalu peka, sehingga uji normalitas tidak perlu dilakukan. Persyaratan yang perlu dipenuhi adalah homogenitas varians populasi.⁹¹

Berdasarkan pendapat-pendapat para ahli di atas dapat disimpulkan bahwa pada pengujian hipotesis statistik tentang rerata, syarat normalitas data telah terpenuhi bila jumlah data sampel minimal 20.

b. Uji Homogenitas

Uji homogenitas data bertujuan untuk menunjukkan bahwa beberapa kelompok sampel yang akan dianalisis menggunakan statistika parameterik berasal dari populasi yang memiliki varians sama. Dalam penelitian ini uji homogenitas dilakukan menggunakan Program SPSS 22.0 melalui **One-Way ANOVA**. Adapun langkah-langkah pengujian homogenitas menggunakan Program SPSS 22.0 adalah sebagai berikut.⁹²

⁹¹ Dali S. Naga, *Probabilitas dan Sekor pada Hipotesis Statistika* (Jakarta: UPT Penerbitan Universitas Tarumanagara, 2008), hh.98-99.

⁹² Kadir, *op. cit*, hh. 169-170.

- (a) Masukkan data pada **Data View**, pada kolom pertama diberi nama “Bias_Butir” yang memuat banyaknya butir bias pada masing-masing metode *DIF*, dan pada kolom kedua diberi nama “Metode”. Pada **Variable View**, klik **Values** lalu masukkan kode 1 untuk UBTS, kode 2 untuk Khi-Kuadrat Lord, dan kode 3 untuk DSE;
- (b) Buka menu utama **Analyze**, lalu klik **Compare Mean**, pilih **One_Way ANOVA**;
- (c) Pada kotak dialog **One_Way ANOVA** pindahkan varabel “Bias_Butir” ke kotak **Dependent List** dan Variabel “Metode” ke kotak **Factor**;
- (d) Klik **Option**, lalu contreng **Descriptive** dan **Homogeneity of Variance Test**, kemudian klik **Continue**;
- (e) Klik **Post Hoc**, lalu contreng **Tukey** karena jumlah data ketiga kelompok sama yaitu 30, kemudian klik **Continue**;
- (f) Klik **OK**.

Langkah-langkah di atas, sekaligus dapat digunakan untuk menguji hipotesis statistik uji perbedaan rerata menggunakan **One_Way ANOVA**. Dengan demikian dalam sekali analisis, uji homogenitas varians dan uji perbedaan rerata menggunakan **One_Way ANOVA** dapat dilakukan secara bersama-sama. Perbedaannya hanya pada cara membaca tabel **output** SPSS 22.0 yang terkait dengan homogenitas dan uji **One_Way ANOVA**.

Kriteria pengujian:

Pada **output** SPSS 22.0, uji homogenitas varians dibaca pada tabel **Test of Homogeneity of Variances**. Apa bila nilai **Levene Statistic** pada **df1** dan **df2** dengan nilai **sig.<0,05** (signifikan) berarti data kelompok tidak homogen. Sebaliknya, jika nilai **Levene Statistic** pada **df1** dan **df2** dengan nilai **sig.>0,05** (tidak signifikan) berarti data kelompok homogen. Apabila data kelompok terbukti memiliki varians yang sama, maka uji hipotesis menggunakan **One_Way ANOVA** dapat dilanjutkan.

2) Uji Hipotesis

Apabila uji persyaratan analisis telah terpenuhi, maka selanjutnya dapat dilakukan uji hipotesis. Pengujian hipotesis dilakukan melalui 2 tahap sebagai berikut.

1. Tahap pertama, menguji perbedaan rerata jumlah soal bias butir pada ketiga model *DIF* yang digunakan. Analisis yang digunakan adalah **One-Way ANOVA**. Langkah-langkah uji perbedaan rerata menggunakan **One-Way ANOVA** secara manual adalah sebagai berikut.

(1) Menghitung Jumlah Kuadrat Total

$$JK_{tot} = \sum X_{total}^2 - \frac{(\sum X_{total})^2}{N}$$

(2) Menghitung Jumlah Kuadrat Antar Kelompok

$$JK_{antar} = \sum \frac{(\sum X_A)^2}{n_A} - \frac{(\sum X_{total})^2}{N}$$

(3) Menghitung Jumlah Kuadrat Dalam Kelompok

$$JK_{dalam} = JK_{total} - JK_{antar}$$

(4) Menghitung Mean Kuadrat Antar Kelompok

$$MK_{antar} = \frac{JK_{antar}}{a-1}$$

(5) Menghitung Mean Kuadrat Dalam Kelompok

$$MK_{dalam} = \frac{JK_{dalam}}{N-a}$$

(6) Menghitung nilai $F_{hitung} = \frac{MK_{antar}}{MK_{dalam}}$; $dk_{pembilang}=a-1$, dengan a adalah

banyak kelompok sampel, dan $dk_{penyebut}=N-1$, N adalah banyak data seluruh sampel.

(7) Hipotesis yang diuji:

$$H_0 : \mu_A = \mu_B = \mu_C$$

$$H_1 : \text{Bukan } H_0$$

Keterangan:

μ_A : Rerata banyak bias butir pada metode Distribusi Sampling Empiris

μ_B : Rerata banyak bias butir pada metode Khi-Kuadrat Lord

μ_C : Rerata banyak bias butir pada metode Uji Beda Taraf Sukar

Tabel 3.2 Tabel Bantu Anava Satu Jalur

Sumber Variasi	JK	db	MK	F _{hitung}	F _{tabel}	Sig.
JK _{antar}	$\sum \frac{(\sum X_A)^2}{nA} - \frac{(\sum X_{total})^2}{N}$	a-1	$\frac{JK_{antar}}{a-1}$	$\frac{MK_{antar}}{MK_{dalam}}$
JK _{dalam}	$JK_{total} - JK_{antar}$	N-a	$\frac{JK_{dalam}}{N-a}$
JK _{total}	$\sum X_{total}^2 - \frac{(\sum X_{total})^2}{N}$	N-1

Kriteria pengujian:

Jika $F_{hitung} > F_{tabel}$, maka tolak H_0 dan terima H_1 yang berarti bahwa terdapat perbedaan yang signifikan rerata jumlah bias butir pada ketiga metode *DIF*. Sebaliknya, jika $F_{hitung} < F_{tabel}$, maka terima H_0 dan tolak H_1 yang berarti bahwa tidak terdapat perbedaan yang signifikan rerata jumlah bias butir pada ketiga metode *DIF*.

Dalam penelitian ini uji hipotesis dilakukan menggunakan Program SPSS 22.0. Langkah-langkah pengujian hipotesis sama dengan langkah pada pengujian homogenitas di atas (langkah-langkah (a) sampai (f)).

Kriteria pengujian:

Pada **output** SPSS 22.0, uji hipotesis dibaca pada tabel **ANOVA**. Apa bila nilai **F** dengan nilai **sig.<0,05** (signifikan), maka tolak H_0 dan terima H_1

yang berarti bahwa terdapat perbedaan yang signifikan rerata jumlah bias butir pada ketiga metode *DIF*. Sebaliknya, jika nilai F dengan nilai $sig.>0,05$ (tidak signifikan), maka tolak H_0 dan terima H_1 yang berarti bahwa tidak terdapat perbedaan yang signifikan rerata jumlah bias butir pada ketiga metode *DIF*.

2. Tahap kedua, apabila terbukti terdapat perbedaan yang signifikan rerata jumlah bias butir pada ketiga metode *DIF*, maka dilanjutkan uji *Post Hoc*

menggunakan Uji Tukey menggunakan rumus: $Q = \frac{\mu_i - \mu_j}{\sqrt{\frac{MK_{dalam}}{n}}}$; $dk=n-1$

dengan:

Hipotesis yang diuji:

$$(1) \quad H_0 : \mu_A - \mu_B = 0$$

$$H_1 : \mu_A - \mu_B > 0$$

Keterangan:

μ_A : Rerata banyak bias butir pada metode Distribusi Sampling Empiris

μ_B : Rerata banyak bias butir pada metode Khi-Kuadrat Lord

$$(2) \quad H_0 : \mu_B - \mu_C = 0$$

$$H_1 : \mu_B - \mu_C > 0$$

Keterangan:

μ_B : Rerata banyak bias butir pada metode Khi-Kuadrat Lord

μ_C : Rerata banyak bias butir pada metode Uji Beda Taraf Sukar

Uji **Post Hoc** juga dilakukan dengan Program SPSS 22.0, langkah-langkah pengujian sama dengan langkah-langkah uji homogenitas dan uji hipotesis yang telah dipaparkan di atas (langkah-langkah (a) sampai dengan (f)).

Kriteria pengujian:

Uji **Post Hoc** dapat dibaca pada tabel **Multiple Comparisons, output** SPSS 22.0. Perbedaan rerata masing-masing metode *DIF* dapat dibaca pada kolom *Mean Difference (I-J)*. Jika nilai *Mean Difference (I-J)* dengan nilai $\text{sig.}p < 0,05$ (signifikan), berarti bahwa perbedaan rerata banyaknya bias butir pada metode *DIF* ke-i dan metode *DIF* ke-j signifikan. Dengan demikian metode *DIF* ke-i dikatakan lebih sensitif daripada metode *DIF* ke-j. Sebaliknya, jika nilai *Mean Difference (I-J)* dengan nilai $\text{sig.}p > 0,05$ (tidak signifikan), berarti bahwa perbedaan rerata banyaknya bias butir pada metode *DIF* ke-i dan metode *DIF* ke-j tidak signifikan.

G. Hipotesis Statistik

Sesuai dengan rumusan hipotesis penelitian yang telah dipaparkan dalam Bab II disertasi ini, maka hipotesis statistik dapat dirumuskan sebagai berikut.

$$1. \quad H_0 \quad : \quad \mu_A - \mu_B = 0$$

$$H_1 \quad : \quad \mu_A - \mu_B > 0$$

$$2. \quad H_0 \quad : \quad \mu_B - \mu_C = 0$$

$$H_1 \quad : \quad \mu_B - \mu_C > 0$$

Keterangan:

μ_A : Rerata banyak bias butir pada metode Distribusi Sampling Empiris

μ_B : Rerata banyak bias butir pada metode Khi-Kuadrat Lord

μ_C : Rerata banyak bias butir pada metode Uji Beda Taraf Sukar

BAB IV

HASIL PENELITIAN DAN PEMBAHASAN

Secara umum, data hasil penelitian disajikan dalam tiga kelompok besar yaitu: (1) deskripsi data hasil penelitian, (2) analisis data, dan (3) pembahasan hasil penelitian. Untuk mendapatkan gambaran lebih jelas dan lengkap terhadap hasil penelitian, secara rinci data-data tersebut diuraikan dalam masing-masing subbab berikut.

A. Deskripsi Data

Pada bagian berikut dideskripsikan hasil penelitian berturut-turut: (1) data jumlah respons peserta didik yang mengerjakan UN Mata Pelajaran Matematika SMA Program IPA Provinsi DKI Jakarta Tahun Pelajaran 2014/2015 paket kode 1101, (2) indikator pada masing-masing butir pada paket kode 1101, (3) uji kecocokan model, (4) data parameter butir, (5) fungsi informasi butir, dan (6) data hasil penelitian berupa banyak bias butir yang terdeteksi pada masing-masing metode *DIF* yang digunakan.

1. Data Respons Peserta

Pada bagian pertama, disajikan data peserta didik yang memberikan respons terhadap UN Mata Pelajaran Matematika SMA Program IPA Provinsi DKI Jakarta Tahun Pelajaran 2014/2015. Secara keseluruhan, jumlah peserta UN Mata Pelajaran Matematika SMA Negeri dan Swasta Program IPA

Provinsi DKI Jakarta Tahun Pelajaran 2014/2015 sebanyak 23.519 orang yang tersebar pada 397 sekolah. Dari jumlah 23.519 orang peserta didik tersebut, dalam pelaksanaan UN tahun pelajaran 2014/2015 mengerjakan paket soal yang berbeda-beda. Khusus untuk Mata Pelajaran Matematika SMA Program IPA dengan kode paket 1101, dikerjakan oleh 4.869 orang dengan rincian sebagai berikut.

Tabel 4.1 Data Peserta UN Matematika SMA-IPA Paket 1101

No.	Peserta Didik	Jumlah
1.	Laki-laki	2.114
2.	Perempuan	2.755
	Jumlah	4.869

2. Kompetensi dan Indikator Kisi-Kisi UN 2014/2015

Selanjutnya disajikan data yang berkaitan dengan kompetensi dan indikator yang diujikan pada masing-masing butir soal UN. Materi uji (kompetensi) pada kisi-kisi UN tahun pelajaran 2014/2015 mengacu pada Standar Kompetensi Lulusan (SKL) masing-masing mata pelajaran. SKL mata pelajaran dijabarkan dalam bentuk Standar Kompetensi (SK) dan Kompetensi Dasar (KD), sebagaimana yang tercantum dalam Peraturan Menteri Pendidikan Nasional (Permendiknas) No. 22 Tahun 2006 tentang

Standar Isi. Indikator dalam kisi-kisi UN Tahun Pelajaran 2014/2015 diturunkan dari kompetensi yang dirangkum dari beberapa KD.

Tabel 4.2 Indikator pada Kisi-Kisi UN 2014/2015

No. Butir	Indikator
1.	Menentukan penarikan kesimpulan dari beberapa premis.
2.	Menentukan ingkaran atau kesetaraan dari pernyataan majemuk atau pernyataan berkuantor.
3.	Menggunakan aturan pangkat, akar, dan logaritma.
4.	Menggunakan aturan pangkat, akar, dan logaritma.
5.	Menggunakan aturan pangkat, akar, dan logaritma.
6.	Menggunakan rumus jumlah dan hasil kali akar-akar persamaan kuadrat.
7.	Menyelesaikan masalah persamaan atau fungsi kuadrat dengan menggunakan diskriminan.
8.	Menyelesaikan masalah sehari-hari yang berkaitan dengan sistem persamaan linear.
9.	Menentukan persamaan lingkaran atau garis singgunglingkaran.
10.	Menentukan persamaan lingkaran atau garis singgung lingkaran.
11.	Menyelesaikan masalah yang berkaitan dengan teorema sisa atau teorema faktor.
12.	Menyelesaikan masalah yang berkaitan dengan teorema sisa atau teorema faktor.
13.	Menyelesaikan masalah yang berkaitan dengan komposisi dua fungsi atau fungsi invers.
14.	Menyelesaikan masalah program linear.

No. Butir	Indikator
15.	Menyelesaikan operasi matriks.
16.	Menyelesaikan operasi aljabar beberapa vektor dengan syarat tertentu.
17.	Menyelesaikan operasi aljabar beberapa vektor dengan syarat tertentu.
18.	Menyelesaikan masalah yang berkaitan dengan besar sudut atau nilai perbandingan trigonometri sudut antara dua vektor.
19.	Menentukan bayangan titik atau kurva karena dua transformasi atau lebih.
20.	Menyelesaikan masalah yang berkaitan dengan fungsi eksponen atau fungsi logaritma.
21.	Menyelesaikan masalah yang berkaitan dengan fungsi eksponen atau fungsi logaritma.
22.	Menyelesaikan masalah deret aritmetika.
23.	Menyelesaikan masalah deret geometri.
24.	Menghitung jarak dan sudut antara dua objek (titik, garis dan bidang) di ruang dimensi tiga.
25.	Menyelesaikan masalah geometri dengan menggunakan aturan sinus atau kosinus.
26.	Menyelesaikan masalah geometri dengan menggunakan aturan sinus atau kosinus.
27.	Menyelesaikan persamaan trigonometri.
28.	Menyelesaikan masalah yang berkaitan dengan nilai perbandingan trigonometri yang menggunakan rumus jumlah dan selisih sinus, kosinus dan tangen serta jumlah dan selisih dua sudut.

No. Butir	Indikator
29.	Menghitung nilai limit fungsi aljabar dan fungsi Trigonometri.
30.	Menghitung nilai limit fungsi aljabar dan fungsi Trigonometri.
31.	Menyelesaikan soal aplikasi turunan fungsi.
32.	Menentukan integral tak tentu dan integral tentu fungsi aljabar dan fungsi trigonometri.
33.	Menentukan integral tak tentu dan integral tentu fungsi aljabar dan fungsi trigonometri.
34.	Menentukan integral tak tentu dan integral tentu fungsi aljabar dan fungsi trigonometri.
35.	Menentukan integral tak tentu dan integral tentu fungsi aljabar dan fungsi trigonometri.
36.	Menghitung luas daerah dan volume benda putar dengan menggunakan integral.
37.	Menghitung luas daerah dan volume benda putar dengan menggunakan integral.
38.	Menghitung ukuran pemusatan atau ukuran letak dari data dalam bentuk tabel, diagram atau grafik.
39.	Menyelesaikan masalah sehari-hari dengan menggunakan kaidah pencacahan, permutasi, atau kombinasi.
40.	Menyelesaikan masalah yang berkaitan dengan peluang suatu kejadian.

3. Pencocokan Model

Sebagaimana telah diuraikan sebelumnya bahwa pencocokan model merupakan salah satu keputusan penting dalam *IRT*. Semua kegiatan perhitungan dilakukan sesuai dengan model. Bahkan invariansi pada parameter butir dan kemampuan peserta tes ditentukan berdasarkan model tersebut. Pencocokan model dilakukan dengan Program BILOG 3.02. Kecocokan butir dengan model logistik tiga parameter (L3P) dapat dilihat dari harga *p-value* pada kolom *CHISQ output.PH2* yang dijalankan dengan program BILOG 3.02. Butir-butir yang memiliki harga *p-value* lebih besar dari 0,05 dinyatakan cocok dengan model logistik tiga parameter (L3P).

Untuk menguji kecocokan model digunakan sampel sebanyak 700 data respons peserta didik yang dipilih dengan teknik random sampling. Pemilihan jumlah 700 peserta didik karena menggunakan model L3P, dan jumlah sampel telah cukup setelah diuji kecukupan sampel menggunakan Program SPSS 22.0, sebagai berikut.

Tabel 4.3 Uji Kecukupan Sampel

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,940
Bartlett's Test of Sphericity	Approx. Chi-Square	8816,630
	df	780
	Sig.	,000

Data pada Tabel 4.3 menyatakan bahwa nilai *Kaiser-Meyer-Olkin Measure of Sampling Adequacy* sebesar 0,940 lebih besar dari 0,5 dengan sig. $0,000 < 0,05$ (signifikan) yang berarti bahwa jumlah sampel 700 orang telah cukup. Uji kecocokan model tidak menggunakan replikasi. Hasil uji kecocokan model selengkapnya adalah sebagai berikut.

Tabel 4.4 Pencocokan Model

No. Butir	CHISQ	Probabilitas	Keterangan
1.	3,1	0,7954	Cocok
2.	47,2	0,0000	Tidak Cocok
3.	19,0	0,0043	Tidak Cocok
4.	9,0	0,1722	Cocok
5.	8,2	0,1462	Cocok
6.	3,6	0,7315	Cocok
7.	3,5	0,7522	Cocok
8.	34,1	0,0000	Tidak Cocok
9.	3,1	0,7952	Cocok
10.	6,8	0,2326	Cocok
11.	2,6	0,8559	Cocok
12.	95,5	0,0000	Tidak Cocok
13.	79,3	0,0000	Tidak Cocok
14.	10,6	0,2241	Cocok
15.	6,9	0,2250	Cocok
16.	5,1	0,4062	Cocok
17.	4,8	0,4407	Cocok
18.	17,7	0,0071	Tidak Cocok
19.	9,3	0,1540	Cocok

No. Butir	CHISQ	Probabilitas	Keterangan
20.	3,8	0,7040	Cocok
21.	3,6	0,6078	Cocok
22.	6,1	0,4165	Cocok
23.	12,8	0,0452	Tidak Cocok
24.	38,5	0,0000	Tidak Cocok
25.	9,4	0,1527	Cocok
26.	4,4	0,4892	Cocok
27.	5,6	0,3499	Cocok
28.	11,8	0,0672	Cocok
29.	3,0	0,8851	Cocok
30.	4,2	0,6466	Cocok
31.	28,8	0,0002	Tidak Cocok
32.	19,4	0,0017	Tidak Cocok
33.	8,6	0,1993	Cocok
34.	10,6	0,0586	Cocok
35.	3,3	0,6505	Cocok
36.	1,9	0,8669	Cocok
37.	4,3	0,6367	Cocok
38.	5,2	0,5181	Cocok
39.	9,8	0,2020	Cocok
40.	14,9	0,0367	Tidak Cocok

Setelah dilakukan uji kecocokan model menggunakan Program BILOG 3.02, dari 40 butir yang diuji, hanya 29 butir yang cocok dengan model L3P, sedangkan 11 butir lainnya tidak cocok dengan model. Apabila butir-butir tersebut telah cocok dengan model, berarti syarat-syarat unidimensi,

independensi lokal, dan invarian parameter telah terpenuhi untuk masing-masing butir tersebut. Dengan demikian, hanya butir-butir yang cocok dengan model yang dapat dianalisis menggunakan *IRT*, sehingga hanya 29 butir tersebut yang akan dilibatkan dalam analisis berikutnya.

Untuk memudahkan perhitungan dalam analisis selanjutnya, butir yang cocok dengan model diberi nomor kode 1 sampai dengan 29. Nomor kode dan nomor butir selengkapnya dapat dilihat dalam tabel berikut.

Tabel 4.5 Kode Butir yang Cocok Model

No. Kode	No. Butir	No. Kode	No. Butir	No. Kode	No. Butir
1.	1	11.	16	21.	29
2.	4	12.	17	22.	30
3.	5	13.	19	23.	33
4.	6	14.	20	24.	34
5.	7	15.	21	25.	35
6.	9	16.	22	26.	36
7.	10	17.	25	27.	37
8.	11	18.	26	28.	38
9.	14	19.	27	29.	39
10.	15	20.	28		

4. Parameter Butir

Parameter butir diestimasi menggunakan Program BILOG 3.02, yang dapat dibaca pada *output.PH2*. Dalam penelitian ini menggunakan model L3P, sehingga terdapat 3 parameter butir yang diestimasi yaitu parameter daya pembeda (a), taraf sukar (b), dan kebetulan menjawab betul (c).

Dengan adanya parameter kebetulan menjawab betul (c) pada model logistik tiga parameter, memungkinkan peserta tes dengan kemampuan rendah mempunyai peluang lebih besar untuk menjawab dengan benar butir soal tertentu. Berikut ini dipaparkan hasil analisis parameter butir soal UN mata pelajaran Matematika kode paket 1101 untuk Provinsi DKI Jakarta tahun pelajaran 2014/2015.

Tabel 4.6 Parameter Butir Soal UN Paket 1101

No. Butir	Parameter		
	a	b	c
1.	1.122	0.262	0.209
2.	0.347	-0.084	0.141
3.	0.911	-1.463	0.168
4.	0.913	-1.436	0.148
5.	1.520	-0.344	0.129
6.	1.070	-0.665	0.211
7.	0.936	-0.107	0.132
8.	0.318	-0.473	0.172
9.	2.114	0.167	0.252
10.	1.338	-0.422	0.078
11.	1.584	0.076	0.367
12.	0.230	0.881	0.165
13.	0.139	-3.059	0.224
14.	1.169	2.872	0.394
15.	0.947	-1.835	0.171
16.	1.116	-1.608	0.157

No. Butir	Parameter		
	a	b	c
17.	1.876	-0.013	0.112
18.	1.091	-0.903	0.101
19.	2.347	0.362	0.252
20.	1.335	0.404	0.179
21.	2.313	0.262	0.189
22.	1.171	-0.958	0.223
23.	1.170	-0.409	0.309
24.	1.012	0.098	0.082
25.	2.100	0.444	0.094
26.	1.244	-0.800	0.167
27.	2.365	0.208	0.200
28.	1.122	-0.211	0.082
29.	1.467	0.224	0.322
30.	1.221	0.003	0.203
31.	0.782	-0.086	0.215
32.	1.217	-0.936	0.094

No.	Parameter		
Butir	a	b	c
33.	2.346	0.113	0.256
34.	2.428	0.074	0.177
35.	2.477	0.263	0.160
36.	1.440	0.228	0.148

No.	Parameter		
Butir	a	b	c
37.	2.422	0.315	0.252
38.	0.769	-1.388	0.181
39.	1.365	0.092	0.272
40.	0.797	-0.230	0.255

Secara praktis nilai parameter daya pembeda (a) berkisar antara 0 sampai 2. Pada Tabel 4.6, dapat dideskripsikan bahwa parameter daya pembeda (a) memiliki nilai tertinggi 2,477 pada butir soal nomor 35 dan terendah 0,139 pada butir soal nomor 13. Pada butir soal yang baik, nilai a mempunyai hubungan yang positif antara parameter butir dengan kemampuan peserta tes. Dengan demikian maka butir soal nomor 35, menyatakan butir soal yang paling peka terhadap perbedaan kemampuan peserta tes. Sedangkan butir soal nomor 13 merupakan butir soal yang paling tidak peka. Hasil analisis ini juga sesuai dengan uji kecocokan model, yang menyatakan butir soal nomor 13 tidak cocok dengan model. Kepekaan ini tampak pada nilai probabilitas jawaban betul. Pada butir yang peka, sedikit saja kemampuan responden berbeda mengakibatkan cukup besar perbedaan nilai probabilitas jawaban betul.

Parameter taraf sukar (b) secara teoretik nilainya terletak di antara $-\infty$ dan $+\infty$, tetapi secara praktis berkisar antara -3 dan 3. Pengkategorian taraf

sukar butir, sebagaimana dikemukakan oleh Effendi (2011) adalah sebagai berikut.

Tabel 4.7 Kategori Taraf Sukar

No.	Interval	Kategori
1.	$-3,00 \leq b \leq -2,00$	Sangat Mudah
2.	$-1,99 \leq b \leq -1,00$	Mudah
3.	$-0,99 \leq b \leq 1,00$	Sedang
4.	$1,01 \leq b \leq 2,00$	Sukar
5.	$2,01 \leq b \leq 3,00$	Sangat Sukar

Sesuai dengan analisis pada Tabel 4.7 terlihat bahwa nilai b tertinggi adalah 2,872 terdapat pada butir soal nomor 14. Hal ini berarti bahwa butir soal nomor 14 merupakan butir soal yang paling sulit. Sedangkan nilai taraf sukar terendah sebesar -3,059 terdapat pada butir soal nomor 13, merupakan soal yang paling mudah.

Paramater kebetulan menjawab betul (c) merupakan asimtot bawah yang tidak nol (*non zero lower asymptot*) pada kurva karakteristik butir. Sebuah butir soal dikatakan baik, bila nilai c tidak lebih dari $1/k$, dengan k adalah banyaknya alternatif jawaban yang disediakan. Pada tabel 4.6 di atas terlihat bahwa nilai c tertinggi sebesar 0,394 terdapat pada butir soal nomor 14 dan terendah sebesar 0,078 terdapat pada butir soal nomor 10. Soal UN

Matematika SMA tahun pelajaran 2014/2015 terdiri dari 5 pilihan jawaban, berarti soal yang baik memiliki nilai $c \leq 0,2$. Berdasarkan hasil analisis di atas terdapat 26 butir soal dengan nilai $c \leq 0,2$, dan 14 butir soal dengan nilai $c > 0,2$.

5. Nilai Fungsi Informasi Butir

Nilai fungsi informasi butir dapat menjelaskan kekuatan suatu butir dan menyatakan kontribusi butir untuk mengungkapkan kemampuan laten (*latent traits*) yang diukur menggunakan butir soal tersebut. Nilai fungsi informasi butir maksimum dapat dibaca pada *output.PH3* hasil analisis Program BILOG 3.02. Secara grafis, nilai fungsi informasi dapat ditemukan melalui kecuraman lengkungan. Semakin curam lengkungan, semakin besar nilai fungsi informasinya. Berikut ini akan dideskripsikan nilai fungsi informasi butir maksimum untuk masing-masing butir.

Tabel 4.8 Nilai Fungsi Informasi Butir Maksimum

No. Butir	$I(\theta)_{maks}$	No. Butir	$I(\theta)_{maks}$	No. Butir	$I(\theta)_{maks}$
1.	0.6097	11.	0.8849	21.	2.689
2.	0.0663	12.	0.028	22.	0.6456
3.	0.4349	13.	0.0091	23.	0.5449
4.	0.4528	14.	0.4555	24.	0.631
5.	1.3024	15.	0.4665	25.	2.653
6.	0.5516	16.	0.6661	26.	0.8105
7.	0.4907	17.	2.0463	27.	2.7525
8.	0.0527	18.	0.7062	28.	0.7749
9.	1.9889	19.	2.4537	29.	0.8334
10.	1.1108	20.	0.9138	30.	0.7306

No. Butir	$I(\theta)_{maks}$
31.	0.2921
32.	0.8921
33.	2.4325
34.	3.0342

No. Butir	$I(\theta)_{maks}$
35.	3.2595
36.	1.1289
37.	2.6134
38.	0.3022

No. Butir	$I(\theta)_{maks}$
39.	0.7982
40.	0.2815

Pada Tabel 4.8 terlihat bahwa nilai fungsi informasi butir maksimum tertinggi sebesar 3,2595 terdapat pada butir soal nomor 35 dan terendah sebesar 0,0091 terdapat pada butir soal nomor 13. Hal ini berarti bahwa butir soal nomor 35 dapat memberikan informasi yang baik dan berkontribusi tinggi untuk mengungkapkan kemampuan laten (*latent traits*) yang diukur menggunakan butir soal tersebut. Sebaliknya butir soal nomor 13 memberikan informasi yang rendah serta kurang berkontribusi untuk mengungkapkan kemampuan laten (*latent traits*) yang diukur menggunakan butir soal tersebut. Hasil analisis ini juga sesuai dengan uji kecocokan model, sehingga butir soal nomor 13 tidak termasuk dalam analisis butir selanjutnya.

6. Data Bias Butir Hasil Deteksi Metode *DIF*

Untuk melakukan analisis bias butir menggunakan tiga metode *DIF* yaitu: (a) metode Uji Beda Taraf Sukar (UBTS), (b) Khi-Kuadrat Lord, dan (c) Distribusi Sampling Empiris (DSE). Analisis dilakukan dengan 2 tahap, tahap pertama analisis parameter butir, matriks varians-kovarians, dan parameter kemampuan (θ) untuk masing-masing kelompok referensi dan kelompok fokal serta subkelompoknya dianalisis menggunakan Program BILOG-MG 3.07.

Tahap kedua, perhitungan analisis bias butir untuk ketiga metode *DIF* yang diperbandingkan menggunakan Program Excel 2010. Untuk mendeteksi adanya bias butir pada masing-masing metode *DIF* dilakukan sebanyak 30 replikasi. Hasil selengkapnya dapat dilihat pada tabel 4.9 berikut.

Tabel 4.9 Bias Butir yang Terdeteksi

Replikasi ke:	Nomor Kode Bias Butir Berdasarkan Metode <i>DIF</i>		
	Uji Beda Taraf Sukar	Khi-Kuadrat Lord	Distribusi Sampling Empiris
1.	1, 13, 24	1, 13, 19, 23, 24	1, 4, 11, 13, 14, 19, 22, 23, 24, 28
2.	25	12, 14, 18, 25	11, 12, 13, 16, 25, 26
3.	1, 8	13, 18, 23	1, 13, 16, 18, 21, 23, 28
4.	1, 7, 14	1, 2, 7, 11, 14	1, 6, 7, 9, 14, 23, 28
5.	14	19, 23, 28	7, 9, 11, 14, 28
6.	1	1, 4, 13, 15, 19, 27	1, 4, 7, 11, 17, 22, 24, 27
7.	1, 13	13, 18, 23	2, 7, 12, 21, 22
8.	1, 13, 14, 15, 17, 28	1, 11, 13, 14, 15, 19, 23, 27, 28	3, 6, 9, 11, 13, 14, 19, 21, 23, 27, 28
9.	6, 8, 12, 14, 22	2, 6, 14, 22, 23	2, 7, 9, 14, 19, 22, 28
10.	1, 11, 13, 14, 15, 17, 23	1, 11, 13, 14, 15, 23	1, 2, 9, 11, 13, 14, 16, 23, 28
11.	1	2, 7, 12	2, 3, 4, 8, 9, 15, 20, 26
12.	3	3, 23	2, 3, 10
13.	1, 21	6, 16, 24	7, 9, 16, 24
14.	1, 11, 13, 14, 17	1, 11, 13, 14, 17	1, 9, 11, 13, 16, 24, 28

Replikasi ke:	Nomor Kode Bias Butir Berdasarkan Metode <i>DIF</i>		
	Uji Beda Taraf Sukar	Khi-Kuadrat Lord	Distribusi Sampling Empiris
15.	14	6, 7, 8, 12, 13, 14, 16, 17	5, 6, 7, 8, 9, 12, 13, 14, 17, 21
16.	1, 8, 14, 15	1, 4, 14, 15, 28	1, 9, 13, 15, 18, 19, 22, 28
17.	12	2, 4, 18	1, 2, 7, 9, 12, 13, 18, 22, 23, 25, 28
18.	1, 23, 26	7, 9, 25	2, 5, 9, 25, 29
19.	21, 25	2, 7, 8, 12, 14, 17, 21, 25, 29	1, 2, 4, 7, 8, 9, 12, 14, 15, 17, 21, 25, 26
20.	1	2, 18, 29	1, 2, 9, 12, 18, 20, 29
21.	1, 6	1, 2, 6, 7, 11, 13, 29	2, 3, 11, 13, 27, 28, 29
22.	8, 15	13, 15	2, 11, 15, 16, 28
23.	21	5, 7, 21, 25	5, 7, 12, 15, 16, 21, 24, 25, 28
24.	7	3, 4, 7, 12, 16, 17	2, 3, 7, 8, 10, 13, 16, 20, 21, 28, 29
25.	17	13, 17	2, 3, 9, 18, 24, 29
26.	1, 11, 13, 14	1, 11, 13, 19, 27, 28	4, 6, 7, 9, 11, 13, 16, 17, 19, 24, 26, 27, 28, 29
27.	1, 10, 11, 13, 14, 15, 17, 19, 24, 29	1, 11, 13, 14, 15, 17, 19, 22, 24	1, 11, 13, 15, 18, 19, 22, 24, 26, 28
28.	1	2, 7, 12, 13, 16, 17, 20, 21, 24, 25, 29	2, 5, 7, 8, 9, 13, 16, 20, 21, 22, 25, 28, 29
29.	1, 21	1, 2, 7, 13, 24	1, 2, 4, 5, 9, 16, 24, 28

Replikasi ke:	Nomor Kode Bias Butir Berdasarkan Metode <i>DIF</i>		
	Uji Beda Taraf Sukar	Khi-Kuadrat Lord	Distribusi Sampling Empiris
30.	1, 6, 14	1, 13, 14, 17, 22	1, 4, 7, 9, 11, 12, 13, 14, 15, 16, 22, 28

Data pada Tabel 4.9 menunjukkan nomor kode bias butir yang terdeteksi pada masing-masing metode *DIF* yang diperbandingkan untuk setiap replikasi. Banyaknya butir yang dinyatakan bias bervariasi untuk masing-masing metode *DIF*, pada setiap replikasi. Ringkasan tentang banyak butir yang terdeteksi bias untuk masing-masing metode *DIF* pada tiap-tiap replikasi dapat dilihat dalam tabel 4.10 berikut.

Tabel 4.10 Ringkasan Banyak Bias Butir pada Metode *DIF*

Replikasi ke:	UBTS	LORD	DSE
1.	3	5	10
2.	1	4	6
3.	2	3	7
4.	3	5	7
5.	1	3	5
6.	1	6	8
7.	2	3	5
8.	6	9	11
9.	5	5	7
10.	7	6	9
11.	1	3	8
12.	1	2	3
13.	2	3	4
14.	5	5	7
15.	1	8	10

Replikasi ke:	UBTS	LORD	DSE
16.	4	5	8
17.	1	3	11
18.	3	3	5
19.	2	9	13
20.	1	3	7
21.	2	7	7
22.	2	2	5
23.	1	4	9
24.	1	6	11
25.	1	2	6
26.	4	6	14
27.	10	9	10
28.	1	11	13
29.	2	5	8
30.	3	5	12
Total	79	150	244
Rerata	2.63	5.00	8.20
Maksimum	10	11	14
Minimum	1	2	3
SD	2.157	2.364	2.833

Pada Tabel 4.10 dapat dijelaskan bahwa rerata banyak butir bias untuk metode UBTS adalah 2,63 dengan standar deviasi 2,157. Banyak bias butir maksimum yang terdeteksi menggunakan metode ini adalah 10 butir terjadi pada replikasi ke-27. Sedangkan banyak butir bias minimal adalah 1 terjadi pada replikasi ke: 2, 5, 6, 11, 12, 15, 17, 20, 23, 24, 25, dan 28. Sebagian besar hasil analisis bias butir menggunakan metode UBTS adalah 1, yang terjadi pada 12 kali replikasi.

Pada metode Khi-Kuadrat Lord, rerata banyak butir bias adalah 5,00 dengan standar deviasi 2,364. Banyak bias butir maksimum yang terdeteksi menggunakan metode ini adalah 11 butir terjadi pada replikasi ke-28. Sedangkan banyak butir bias minimal adalah 2 terjadi pada replikasi ke: 12, 22, dan 25.

Pada metode Distribusi Sampling Empiris (DSE), rerata banyak butir bias adalah 8,20 dengan standar deviasi 2,833. Banyak bias butir maksimum yang terdeteksi menggunakan metode ini adalah 14 butir terjadi pada replikasi ke-26. Sedangkan banyak butir bias minimal adalah 3 terjadi pada replikasi ke-12.

B. Pengujian Persyaratan Analisis

Uji hipotesis yang digunakan dalam penelitian ini adalah uji **One_Way ANOVA**. Sebelum melakukan uji hipotesis menggunakan statistika parameterik **One_Way ANOVA**, terlebih dahulu dilakukan uji persyaratan analisis. Uji persyaratan analisis yang dilakukan dalam penelitian ini adalah uji normalitas dan uji homogenitas varians.

1. Uji Normalitas

Data sampel untuk kelompok metode UBTS, Khi-Kuadrat Lord, dan DSE masing-masing berjumlah 30 ($n=30$), bersumber dari replikasi yang dilakukan secara acak terhadap data respons peserta UN pada mata

pelajaran Matematika SMA Program IPA di Provinsi DKI Jakarta tahun pelajaran 2014/2015 dengan jumlah keseluruhan 4.869 responden. Sesuai dengan Teorema Limit Sentral yang dikemukakan oleh Schewhart (1988: 53-61) dan Conover (1980: 444) dalam Agung (2006), serta pendapat Wonnacott and Wonnacott dalam Dali S. Naga (2008), yang menyatakan bahwa syarat normalitas sudah terpenuhi bila jumlah sampel minimal 20 ($n \geq 20$). Dengan demikian maka prasyarat normalitas sudah terpenuhi dan uji hipotesis statistik tentang rerata dapat dilanjutkan.

2. Uji Homogenitas

Uji homogenitas dilakukan menggunakan Program SPSS 22.0 melalui **One_Way ANOVA**. Hipotesis yang diuji adalah sebagai berikut.

$$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2$$

H_1 : Bukan H_0 .

Hasil uji homogenitas menggunakan Program SPSS 22.0 melalui **One_Way ANOVA** adalah sebagai berikut.

Tabel 4.11 Uji Homogenitas

Test of Homogeneity of Variances

BIAS_BUTIR

Levene Statistic	df1	df2	Sig.
1.759	2	87	.178

Tabel 4.11 menyatakan bahwa nilai *Levene Statistic* sebesar 1,759 pada $df_1=2$ dan $df_2=87$ dengan signifikansi $0,178 > 0,05$ (tidak signifikan). Oleh karena nilai $\text{sig.}=0,275$ lebih besar daripada $0,05$ terima H_0 dan tolak H_1 yang berarti bahwa ketiga kelompok data memiliki varians yang homogen. Sehingga pengujian hipotesis dapat dilanjutkan.

C. Pengujian Hipotesis

Oleh karena uji persyaratan analisis telah terpenuhi, maka selanjutnya dapat dilakukan uji hipotesis. Uji hipotesis dilakukan dengan menggunakan Program SPSS 22.0. Analisis yang digunakan adalah **One_Way ANOVA**. Pengujian hipotesis dilakukan melalui 2 tahap sebagai berikut.

1. Tahap pertama, menguji perbedaan rerata jumlah soal bias butir pada ketiga model *DIF* yang digunakan. Analisis yang digunakan adalah **One_Way ANOVA**. Hipotesis yang diuji adalah sebagai berikut.

$$H_0 : \mu_A = \mu_B = \mu_C$$

$$H_1 : \text{Bukan } H_0$$

Keterangan:

μ_A : Rerata banyak bias butir pada metode Distribusi Sampling Empiris

μ_B : Rerata banyak bias butir pada metode Khi-Kuadrat Lord

μ_C : Rerata banyak bias butir pada metode Uji Beda Taraf Sukar

Apabila terbukti terdapat perbedaan yang signifikan rerata banyaknya bias butir pada ketiga metode *DIF*, maka dilanjutkan uji *Post Hoc* menggunakan Uji *Tukey*. Hasil uji tahap pertama, perbedaan rerata banyaknya soal bias butir pada ketiga model *DIF* adalah sebagai berikut.

Tabel 4.12 Uji Perbedaan Rerata

ANOVA

BIAS_BUTIR

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	468.289	2	234.144	38.452	.000
Within Groups	529.767	87	6.089		
Total	998.056	89			

Data pada Tabel 4.12 menunjukkan bahwa nilai F sebesar 38,452 dengan signifikansi $\text{sig.} = 0,000 < 0,05$ (signifikan). Kesimpulan: tolak H_0 dan terima H_1 . Hal itu berarti bahwa terdapat perbedaan yang signifikan rerata jumlah butir bias yang dideteksi menggunakan metode UBTS, Khi-Kuadrat Lord, dan Distribusi Sampling Empiris. Dengan demikian, untuk mengetahui perbandingan sensitivitas metode *DIF* dilanjutkan dengan melakukan uji *Post Hoc*.

2. Tahap kedua, uji hipotesis kedua dilanjutkan karena telah terbukti adanya perbedaan yang signifikan rerata jumlah bias butir pada ketiga metode *DIF*. Uji hipotesis tahap kedua merupakan uji *Post Hoc* yang bertujuan untuk membandingkan sensitivitas ketiga metode *DIF* yang digunakan

untuk mendeteksi bias butir. Uji *Post Hoc* yang digunakan adalah Uji *Tukey* karena jumlah masing-masing sampel ketiga kelompok sama yaitu 30. Uji hipotesis tahap kedua ini memberikan gambaran, metode *DIF* mana yang paling sensitif di antara tiga metode *DIF* yang digunakan. Uji *Post Hoc* juga dilakukan menggunakan Program SPSS 22.0. Hipotesis yang diuji adalah sebagai berikut.

$$(1) \quad H_0 : \mu_A - \mu_B = 0$$

$$H_1 : \mu_A - \mu_B > 0$$

Keterangan:

μ_A : Rerata banyak bias butir pada metode Distribusi Sampling Empiris

μ_B : Rerata banyak bias butir pada metode Khi-Kuadrat Lord

$$(2) \quad H_0 : \mu_B - \mu_C = 0$$

$$H_1 : \mu_B - \mu_C > 0$$

Keterangan:

μ_B : Rerata banyak bias butir pada metode Khi-Kuadrat Lord

μ_C : Rerata banyak bias butir pada metode Uji Beda Taraf Sukar

Hasil uji *Post Hoc* menggunakan Program SPSS 22.0, dapat dibaca melalui **output** pada tabel **Multiple Comparisons**. Pada tabel tersebut

diperlihatkan perbedaan rerata masing-masing metode yang diperbandingkan. Hasil selengkapnya disajikan dalam tabel 4.13 sebagai berikut.

Tabel 4.13 Hasil Uji *Post Hoc*

Multiple Comparisons

Dependent Variable: BIAS_BUTIR

Tukey HSD

(I) METODE	(J) METODE	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
UBTS	LORD	-2.367*	.637	.001	-3.89	-.85
	DSE	-5.567*	.637	.000	-7.09	-4.05
LORD	UBTS	2.367*	.637	.001	.85	3.89
	DSE	-3.200*	.637	.000	-4.72	-1.68
DSE	UBTS	5.567*	.637	.000	4.05	7.09
	LORD	3.200*	.637	.000	1.68	4.72

*. The mean difference is significant at the 0.05 level.

Data hasil uji *Post Hoc* pada Tabel 4.13 dapat dijelaskan sebagai berikut.

- 1) Metode Distribusi Sampling Empiris (DSE) dan metode Khi-Kuadrat Lord memiliki perbedaan rerata (*Mean Difference*) sebesar 3,200 dengan nilai sig.0,000<0,05 (signifikan), kesimpulan: tolak H_0 dan terima H_1 . Hal itu berarti bahwa rerata hasil deteksi bias butir menggunakan metode DSE sebesar 8,20 dan rerata hasil deteksi bias butir menggunakan metode Khi-Kuadrat Lord sebesar 5,00 dengan perbedaan rerata sebesar 3,200 berbeda secara signifikan. Dengan demikian dapat dinyatakan bahwa metode DSE lebih sensitif

mendeteksi bias butir dibandingkan dengan metode Khi-Kuadrat Lord. Dengan kata lain dapat dinyatakan bahwa hipotesis pertama teruji.

- 2) Metode Khi-Kuadrat Lord dan metode Uji Beda Taraf Sukar (UBTS) memiliki perbedaan rerata (*Mean Difference*) sebesar 2,367 dengan nilai $\text{sig. } 0,001 < 0,05$ (signifikan), kesimpulan: tolak H_0 dan terima H_1 . Hal itu berarti bahwa rerata hasil deteksi bias butir menggunakan metode Khi-Kuadrat Lord sebesar 5,00 dan rerata jumlah hasil deteksi bias butir menggunakan metode UBTS sebesar 2,63 dengan perbedaan rerata sebesar 2,367 berbeda secara signifikan. Dengan demikian dapat dinyatakan bahwa metode Khi-Kuadrat Lord lebih sensitif mendeteksi bias butir dibandingkan dengan metode UBTS. Dengan kata lain dapat dinyatakan bahwa hipotesis kedua teruji.

D. Pembahasan Hasil Penelitian

Sensitivitas metode *DIF* adalah kemampuan suatu metode *DIF* untuk mendeteksi adanya bias butir dalam suatu perangkat tes. Sensitivitas metode *DIF* dapat dilihat dari banyaknya jumlah butir yang terdeteksi mengandung bias butir oleh metode *DIF*. Hasil penelitian ini menyatakan bahwa: (1) metode Distribusi Sampling Empiris (DSE) lebih sensitif daripada metode Khi-Kuadrat Lord untuk mendeteksi bias butir; dan (2) metode Khi-Kuadrat Lord lebih sensitif daripada metode Uji Beda Taraf Sukar (UBTS) untuk mendeteksi bias butir. Pada penelitian ini terbukti metode DSE paling banyak mendeteksi bias butir.

dapat mendeteksi bias butir dengan rerata 8,20; diikuti oleh metode Khi-Kuadrat Lord dengan rerata 5,00; dan metode UBTS dengan rerata 2,63.

Hasil penelitian menyatakan bahwa kedua hipotesis telah teruji. Dalam 30 kali replikasi, secara umum dapat dilihat bahwa sebagian besar butir-butir yang dinyatakan bias oleh metode UBTS juga dinyatakan bias secara keseluruhan (sempurna) oleh metode Khi-Kuadrat Lord dan DSE. Namun terdapat juga sebagian kecil butir-butir yang dinyatakan bias oleh metode *DIF* yang satu, tetapi tidak terdeteksi bias oleh metode *DIF* lainnya. Misalnya terdapat butir-butir yang dinyatakan bias oleh metode UBTS, tidak terdeteksi oleh metode Khi-Kuadrat Lord, tetapi terdeteksi oleh metode DSE. Atau terdapat pula butir-butir yang dinyatakan bias oleh metode UBTS dan Khi-Kuadrat Lord, tetapi tidak terdeteksi oleh metode DSE. Perbedaan tersebut diakibatkan oleh adanya perbedaan teknik dalam perhitungan analisis bias butir pada masing-masing metode *DIF*, yang akan dibahas pada bagian berikut ini.

Pada replikasi pertama, dari 29 butir yang dianalisis menggunakan ketiga metode *DIF*, butir-butir yang dinyatakan bias menggunakan metode UBTS adalah butir dengan nomor kode 1, 13, dan 24 terdeteksi semua (sempurna) oleh metode Khi-Kuadrat Lord. Bahkan terdapat 2 butir lain yang tidak terdeteksi oleh metode UBTS yaitu butir 19 dan 23 terdeteksi bias oleh metode Khi-Kuadrat Lord. Pada replikasi pertama ini, metode Khi-Kuadrat

Lord mendeteksi 5 butir bias. Adapun butir-butir yang dinyatakan bias pada metode Khi-Kuadrat Lord adalah butir nomor kode 1, 13, 19, 23, dan 24.

Sebagai contoh butir nomor kode 19 tidak terdeteksi sebagai bias butir menggunakan metode UBTS. Pada kelompok referensi dengan nilai estimasi parameter butir $a=2,75535$; $b=0,04781$; dan $c=0,21138$. Sedangkan pada kelompok fokal (yang sudah disetarakan) dengan nilai estimasi parameter butir $a=2,15724$; $b=0,15380$; dan $c=0,22497$. Dengan nilai $S_R^2=0,002$ dan $S_F^2=0,003$, sehingga nilai $\sqrt{S_R^2 + S_F^2}=0,7115$ serta $\Delta\hat{b}=-0,10599$. Wajar tidak terdeteksi oleh metode UBTS sebagai butir yang bias karena nilai penyebut cukup besar yaitu 0,7115 sedangkan nilai pembilang $\Delta\hat{b}=-0,10599$ cukup kecil, sehingga nilai $d=-1,48973 > -1,96$ (tidak bias). Berbeda dengan metode Khi-Kuadrat Lord yang memperhitungkan perbedaan ketiga nilai estimasi parameter butir, perbedaan nilai estimasi parameter a sebesar 0,59811; perbedaan nilai estimasi parameter b sebesar 0,10599; dan perbedaan nilai estimasi parameter c sebesar 0,01359.

Perbedaan nilai estimasi parameter a (daya pembeda) mengakibatkan perbedaan kecuraman kurva karakteristik butir. Semakin besar nilai a, maka kurva karakteristik butir semakin curam, dan semakin kecil nilai a maka kurva karakteristik butir semakin landai. Perbedaan nilai estimasi parameter b, mengakibatkan pergeseran kurva karakteristik butir ke kiri atau ke kanan. Semakin besar nilai b (butir tersebut semakin sulit), maka posisi kurva

karakteristik butir semakin ke kanan. Sedangkan perbedaan nilai estimasi parameter c , mengakibatkan pergeseran kurva karakteristik butir ke atas atau ke bawah. Semakin besar nilai c (tebakan semakin tinggi), maka posisi kurva karakteristik butir semakin ke atas. Berbeda dengan metode UBTS yang mengasumsikan parameter daya pembeda sama dan faktor tebakan diasumsikan nol. Asumsi-asumsi tersebut mengakibatkan perbedaan posisi kurva karakteristik kurva pada kelompok referensi dan fokal tidak terdeteksi dengan baik, karena pergeseran-pergeseran posisi kurva karakteristik butir akibat perbedaan nilai a dan c tidak terdeteksi. Dengan demikian metode UBTS kurang peka untuk mendeteksi bias butir bila dibandingkan dengan metode Khi-Kuadrat Lord.

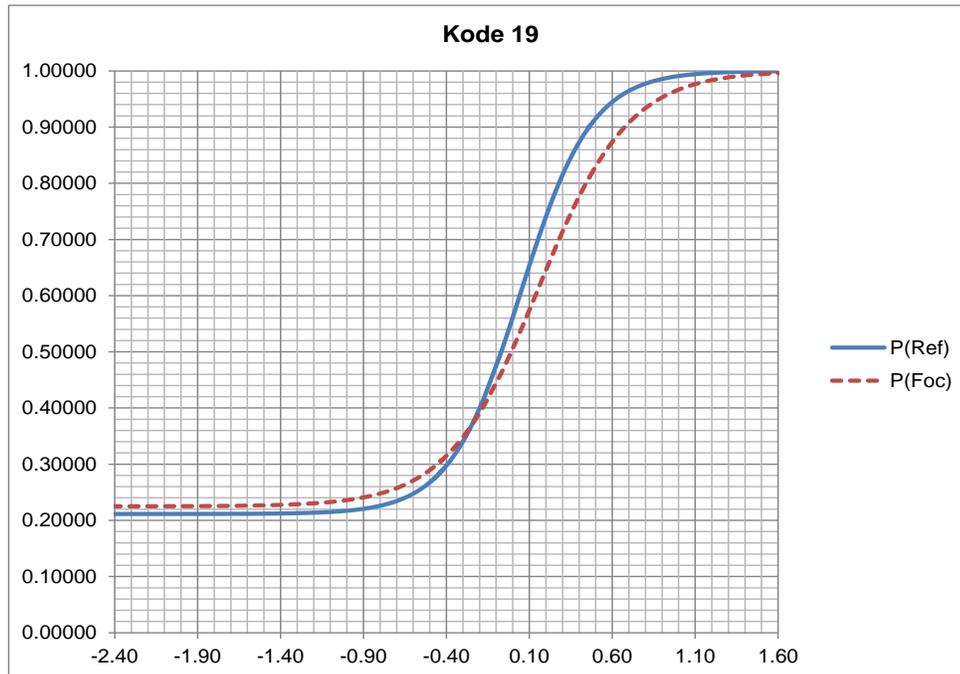
Keunggulan metode Khi-Kuadrat Lord dibandingkan dengan metode UBTS terletak pada kelengkapan parameter butir yang digunakan dalam perhitungan analisis bias butir. Sehingga perbedaan posisi kurva karakteristik butir antara kelompok referensi dan kelompok fokal terlihat secara jelas. Pada butir dengan kode 19, terlihat bahwa kurva karakteristik butir kelompok referensi lebih curam bila dibandingkan dengan kurva karakteristik butir kelompok fokal, karena nilai estimasi parameter daya pembeda kelompok referensi sebesar $a=2,75535$ lebih besar daripada nilai estimasi parameter daya pembeda kelompok fokal (yang sudah disetarakan) sebesar $a=2,15724$. Kurva karakteristik butir kelompok fokal posisinya terlihat di sebelah kanan

kurva karakteristik butir kelompok referensi, karena nilai estimasi parameter butir taraf sukar kelompok fokal sebesar $b=0,15380$ lebih besar daripada nilai estimasi parameter butir taraf sukar kelompok referensi sebesar $b=0,04781$. Asimtot kurva karakteristik butir kelompok fokal terlihat di atas asimtot kurva karakteristik butir kelompok referensi, karena nilai estimasi parameter tebakan semu kelompok fokal sebesar $c=0,22497$ lebih besar daripada nilai estimasi parameter tebakan semu kelompok referensi sebesar $c=0,21138$. Nilai untuk kedua kelompok referensi dan fokal lebih tinggi daripada nilai ideal 0,2 untuk butir soal dengan lima pilihan jawaban. Oleh karena itu termasuk nilai c yang kurang baik. Demikian pula nilai elemen-elemen pada penjumlahan matriks varians kovarians antara kelompok referensi dan fokal

relatif kecil, yaitu $\Sigma = \begin{bmatrix} 0,19642 & 0,01806 & 0,00709 \\ 0,01806 & 0,00499 & 0,00114 \\ 0,00709 & 0,00114 & 0,00135 \end{bmatrix}$, yang menyebabkan nilai

$\chi^2 = 9,73195$ lebih besar daripada $\chi_{kritis}^2 = 7,81472$, sehingga butir dengan kode 19 dinyatakan bias.

Untuk memperjelas uraian di atas, berikut ini disajikan perbandingan kurva karakteristik butir kelompok referensi (laki-laki) dan kelompok fokal (perempuan) soal dengan nomor kode 19.



Gambar 4.1 Kurva karakteristik Butir
Kelompok Referensi dan Fokal Kode 19

Pada replikasi pertama, butir-butir yang dinyatakan bias oleh metode UBTS yaitu butir dengan kode 1, 13, 24 dan butir-butir yang terdeteksi bias oleh metode Khi-Kuadrat Lord pada replikasi pertama 1, 13, 19, 23, dan 24 ternyata terdeteksi seluruhnya (sempurna) juga oleh metode DSE. Bahkan metode DSE mendeteksi 5 butir lain yang dinyatakan tidak bias menurut metode UBTS dan Khi-Kuadrat Lord yaitu butir dengan nomor kode 4, 11, 14, 22, dan 28. Sehingga secara keseluruhan butir-butir yang dinyatakan bias oleh metode DSE pada replikasi pertama adalah butir-butir dengan nomor kode: 1, 4, 11, 13, 14, 19, 22, 23, 24, dan 28.

Keunggulan metode DSE terletak pada perhitungan bias butir menggunakan replikasi analisis. Sebagai contoh butir dengan kode 14, yang tidak terdeteksi bias oleh metode UBTS dan Khi-Kuadrat Lord merupakan butir bias menurut metode DSE. Karakteristik butir nomor kode 14 adalah sebagai berikut: (a) pada kelompok referensi nilai estimasi parameter $a=1,36224$; $b=0,26804$; dan $c=0,17276$; (b) pada kelompok fokal nilai estimasi parameter (sudah disetarakan) $a=1,35090$; $b=0,44861$; dan $c=0,20816$. Selanjutnya nilai *standard error* untuk parameter b (taraf sukar) pada kelompok referensi sebesar nilai $S_R^2=0,005$ dan $S_F^2=0,005$, sehingga nilai $\sqrt{S_R^2 + S_F^2}= 0,10150$ serta $\Delta\hat{b}=-0,18057$. Wajar tidak terdeteksi oleh metode UBTS sebagai butir yang bias karena nilai penyebut dan pembilang memiliki perbedaan nilai absolut yang cukup kecil sehingga nilai $d=-1,77909 > -1,96$ (tidak bias).

Pada metode Khi-Kuadrat Lord, butir dengan kode 14 tersebut memiliki perbedaan nilai estimasi parameter a sebesar 0,01134; perbedaan nilai estimasi parameter b sebesar 0,18057; dan perbedaan nilai estimasi parameter c sebesar 0,03540. Perbedaan nilai estimasi parameter butir yang cukup kecil untuk parameter a dan c, kecuali perbedaan estimasi parameter taraf sukar. Elemen-elemen jumlah matriks varians kovarians kelompok

referensi dan fokal $\Sigma = \begin{bmatrix} 0,08891 & 0,01488 & 0,00889 \\ 0,01488 & 0,01015 & 0,00256 \\ 0,00889 & 0,00256 & 0,00222 \end{bmatrix}$, juga terlihat relatif

kecil, elemen matriks varians nilainya relatif kecil bahkan σ_c^2 mendekati nol berarti sebaran datanya dekat dengan reratanya. Demikian pula elemen matriks kovarians nilainya relatif kecil mendekati nol dan bertanda positif (searah). Oleh karena perbedaan estimasi parameter butir relatif kecil, menyebabkan nilai $\chi^2 = 4,45207$ lebih kecil daripada $\chi_{kritis}^2 = 7,81472$, sehingga butir dengan kode 14 dinyatakan tidak bias menurut metode Khi-Kuadrat Lord.

Perhitungan analisis bias butir menggunakan metode DSE dilakukan dengan menghitung perbedaan probabilitas menjawab benar suatu butir soal ke-i antara kelompok referensi dan fokal. Untuk menghitung probabilitas menjawab benar butir ke-i, diperlukan parameteri butir yang telah dihitung sebelumnya dan parameter kemampuan. Parameter kemampuan (θ) dibangkitkan dari sampel yang dipilih secara acak berjumlah 1000 orang, bersumber dari gabungan kelompok referensi dan fokal secara keseluruhan. Selanjutnya akan diperoleh nilai $P(\theta)$ untuk masing-masing kelompok referensi dan fokal pada masing-masing θ , serta dihitung selisih nilai $\Delta P_i = P_R(\theta_j) - P_F(\theta_j)$. Untuk menghitung indeks bias butir (δ) kelompok referensi dan kelompok fokal menggunakan rumus: $SPD - \theta = \frac{\sum_{j=1}^{n_p} \Delta P(\theta_j)}{n_p}$

untuk bias butir jenis *uniform* (konsisten) dan $UPD - \theta = \sqrt{\frac{\sum_{j=1}^{n_p} (\Delta P(\theta_j))^2}{n_p}}$ untuk

bias butir jenis *non-uniform* (tidak konsisten). Untuk mengetahui apakah butir ke- i merupakan bias butir tipe *uniform* atau *non-uniform*, dapat dilihat dari tanda $\Delta P_i = P_R(\theta_j) - P_F(\theta_j)$. Apabila dalam selang tertentu ΔP_i berbeda tanda (misalnya berubah dari positif menjadi negatif dan sebaliknya), berarti butir tersebut merupakan bias butir jenis *non-uniform*, maka perhitungan indeks

bias butir menggunakan rumus $UPD - \theta = \sqrt{\frac{\sum_{j=1}^{n_p} (\Delta P(\theta_j))^2}{n_p}}$. Sebaliknya apabila

dalam selang tertentu ΔP_i bertanda sama (misalnya positif semua atau negatif semua), berarti butir tersebut merupakan bias butir jenis *uniform*, sehingga perhitungan indeks bias butir menggunakan rumus: $SPD - \theta =$

$\frac{\sum_{j=1}^{n_p} \Delta P(\theta_j)}{n_p}$. Nilai $SPD - \theta$ atau $UPD - \theta$ yang diperoleh disebut sebagai δ .

Selanjutnya dilakukan replikasi analisis terhadap subkelompok referensi (R1-R2) dan focal (F1-F2) setelah masing-masing subkelompok mendapatkan perlakuan yang sama dengan kelompok referensi dan focal sebelumnya. Pada masing-masing subkelompok dihitung $\Delta P_F = P_{F1}(\theta_j) - P_{F2}(\theta_j) = \delta_1$ dan $\Delta P_R = P_{R1}(\theta_j) - P_{R2}(\theta_j) = \delta_2$, serta di tentukan nilai maksimum dari $(\delta_1; \delta_2)$. Bila nilai δ lebih besar daripada nilai maksimum dari $(\delta_1; \delta_2)$, maka butir ke- i dinyatakan sebagai butir bias. Sebaliknya jika nilai δ

lebih kecil daripada nilai maksimum dari $(\delta_1; \delta_2)$, maka butir ke-i dinyatakan sebagai butir tidak bias.

Butir soal dengan kode 14 dinyatakan bias oleh metode DSE. Parameter butir soal dengan kode 14 untuk kelompok referensi dan fokal, sama dengan hasil estimasi parameter butir menggunakan metode Khi-Kuadrat Lord dan UBTS, yaitu: (a) pada kelompok referensi nilai estimasi parameter $a=1,36224$; $b=0,26804$; dan $c=0,17276$; (b) pada kelompok fokal nilai estimasi parameter (sudah disetarakan) $a=1,35090$; $b=0,44861$; dan $c=0,20816$. Memperhatikan adanya perbedaan tanda ΔP_i , yaitu pada interval $-2,418969 \leq \theta \leq -0,55187$; $\Delta P_i < 0$ (negatif), sedangkan pada interval $-0,54906 \leq \theta \leq 1,68553$; $\Delta P_i > 0$ (positif) maka dapat disimpulkan bahwa butir soal kode 14 merupakan bias butir jenis *non-uniform*. Hal itu berarti bahwa pada interval $-2,418969 \leq \theta \leq -0,55187$ butir soal menguntungkan kelompok fokal (perempuan) sedangkan pada interval $-0,54906 \leq \theta \leq 1,68553$ butir soal menguntungkan kelompok referensi (laki-laki). Pada perhitungan analisis bias butir menggunakan rumus $UPD - \theta$ diperoleh nilai $\delta=0,04056$.

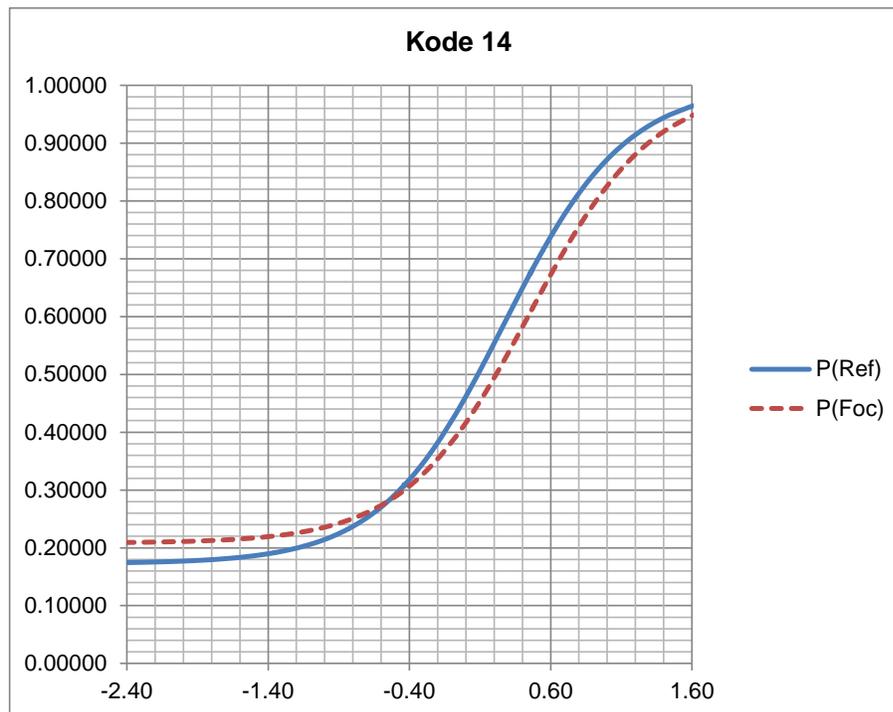
Namun terdapat perbedaan nilai estimasi parameter butir pada subkelompok referensi maupun fokal. Perbedaan tersebut diakibatkan oleh jumlah sampel yang berbeda yaitu setengah kali jumlah sampel kelompok referensi dan fokal, serta perbedaan pola respons peserta yang dianalisis.

Parameter butir untuk subkelompok referensi, R1: $a=1,64655$; $b=0,29942$; dan $c=0,21259$; R2 (sudah disetarakan): $a=1,41011$; $b=0,11730$; $c=0,14973$. Memperhatikan adanya perbedaan tanda ΔP_i , yaitu pada interval $-2,34996 \leq \theta \leq -0,68668$; $\Delta P_i > 0$ (positif), pada interval $-0,66325 \leq \theta \leq 1,14722$; $\Delta P_i < 0$ (negatif), sedangkan pada interval $1,24694 \leq \theta \leq 1,62146$; $\Delta P_i > 0$ (positif), maka dapat disimpulkan bahwa butir soal kode 14 merupakan bias butir jenis *non-uniform*. Hal itu berarti bahwa pada interval $-2,34996 \leq \theta \leq -0,68668$ butir soal menguntungkan kelompok referensi R1, sedangkan pada interval $-0,66325 \leq \theta \leq 1,14722$ butir soal menguntungkan subkelompok referensi R2, dan pada interval $1,24694 \leq \theta \leq 1,62146$ menguntungkan subkelompok referensi R1. Pada perhitungan analisis bias butir menggunakan rumus $UPD - \theta$ diperoleh nilai $\delta_1=0,03617$.

Sedangkan parameter butir untuk subkelompok fokal, F1: $a=1,11890$; $b=0,34553$; dan $c=0,16153$; F2 (sudah disetarakan): $a=1,06737$; $b=0,45586$; dan $c=0,17349$. Memperhatikan adanya perbedaan tanda ΔP_i , yaitu pada interval $-2,27806 \leq \theta \leq -0,64266$; $\Delta P_i < 0$ (negatif), sedangkan pada interval $-0,63567 \leq \theta \leq 1,63787$; $\Delta P_i > 0$ (positif), maka dapat disimpulkan bahwa butir soal kode 14 merupakan bias butir jenis *non-uniform*. Hal itu berarti bahwa pada interval $-2,27806 \leq \theta \leq -0,64266$ butir soal menguntungkan kelompok fokal F2, sedangkan pada interval $-0,63567 \leq \theta \leq 1,63787$ butir soal menguntungkan subkelompok fokal F1. Pada

perhitungan analisis bias butir menggunakan rumus $UPD - \theta$ diperoleh nilai $\delta_2=0,02487$.

Berdasarkan hasil analisis bias butir pada kelompok referensi dan fokal ($\delta=R-F$), serta hasil replikasi analisis pada subkelompok referensi ($\delta_1=R1-R2$) dan fokal ($\delta_2=F1-F2$) diperoleh bahwa nilai $\delta=0,04056$ ternyata lebih besar daripada nilai maksimum $(\delta_1, \delta_2)= 0,03617$ sehingga butir soal dengan kode 14 dinyatakan sebagai butir yang bias menurut metode DSE. Untuk memperoleh gambaran yang lebih jelas tentang butir dengan kode 14, dapat dilihat pada gambar kurva karakteristik butir pada gambar 4.2 berikut.



Gambar 4.2 Kurva Karakteristik Butir Kelompok Referensi dan Fokal Kode 14

Pada butir dengan kode 14 terlihat bahwa kurva karakteristik butir kelompok referensi lebih curam bila dibandingkan dengan kurva karakteristik butir kelompok fokal, karena nilai estimasi parameter daya pembeda kelompok referensi sebesar $a=1,36224$ lebih besar daripada nilai estimasi parameter daya pembeda kelompok fokal (yang sudah disetarakan) sebesar $a=1,35090$. Kurva karakteristik butir kelompok fokal posisinya terlihat di sebelah kanan kurva karakteristik butir kelompok referensi, karena nilai estimasi parameter butir taraf sukar kelompok fokal sebesar $b=0,44861$ lebih besar daripada nilai estimasi parameter butir taraf sukar kelompok referensi sebesar $b=0,26804$. Asimtot kurva karakteristik butir kelompok fokal terlihat di atas asimtot kurva karakteristik butir kelompok referensi, karena nilai estimasi parameter tebakan semu kelompok fokal sebesar $c=0,20816$ lebih besar daripada nilai estimasi parameter tebakan semu kelompok referensi sebesar $c=0,17276$.

Pada pembahasan kedua, diuraikan karakteristik butir yang terdeteksi oleh ketiga metode *DIF* pada replikasi pertama. Butir soal yang terdeteksi oleh ketiga metode *DIF* adalah butir soal dengan kode 1, 13, dan 24. Perhatikan soal dengan nomor kode 1 atau butir soal nomor 1, pada replikasi ke-1. Butir soal tersebut terdeteksi bias butir menggunakan ketiga metode. Pada kelompok referensi, butir tersebut memiliki parameter $a=1,07353$; $b=-0,12977$; $c=0,21498$; dan *standard error* parameter b adalah $S_R=0,123$. Pada

kelompok focal sebelum disetarakan, butir tersebut memiliki parameter $a=1,41973$; $b=0,23435$; $c=0,26926$; dan *standard error* parameter b $S_F=0,087$. Setelah disetarakan menggunakan metode rerata dan sigma, butir tersebut memiliki parameter $a=1,28231$; $b=0,24082$; dan $c=0,26926$. Besarnya konstanta penyetaraan $A=1,10717$; $K=-0,01864$; $\mu_{bR}=-0,20029$; $\mu_{bF}=-0,16406$; $\sigma_{bR}=1,11269$; dan $\sigma_{bF}=1,00499$. Bila dilihat dari parameter daya pembeda, baik untuk kelompok referensi maupun focal, butir soal nomor 1 memiliki daya pembeda yang sangat baik $a \geq 1$. Ditinjau dari parameter taraf sukar, butir soal nomor 1 termasuk soal kategori sedang dengan taraf sukar $-0,99 \leq b \leq 1$. Namun soal ini memiliki nilai kebetulan menjawab betul tidak baik karena nilai c sebesar $0,26926 > 0,200$ untuk 5 pilihan jawaban.

Berikut ini disajikan indikator, rumusan butir soal, dan pembahasan butir soal nomor 1.

- Indikator : Menentukan penarikan kesimpulan dari beberapa premis.
- Soal : Diketahui:
 Premis 1: Ayah tidak ke rumah sakit atau ayah periksa ke dokter.
 Premis 2: Ayah tidak periksa ke dokter.
 Kesimpulan yang sah dari premis-premis tersebut adalah....
 A. Ayah ke rumah sakit.
 B. Ayah tidak ke rumah sakit.

- C. Ayah di rumah saja.
- D. Ayah tidak ke rumah sakit, dan ayah tidak periksa ke dokter.
- E. Ayah ke rumah sakit dan Ayah tidak periksa ke dokter.

Pembahasan : Pernyataan pada premis 1: $\sim p \vee q \equiv p \rightarrow q$, sehingga premis 1 dan 2 berbentuk:

$$\frac{p \rightarrow q}{\sim q} \therefore \sim p$$

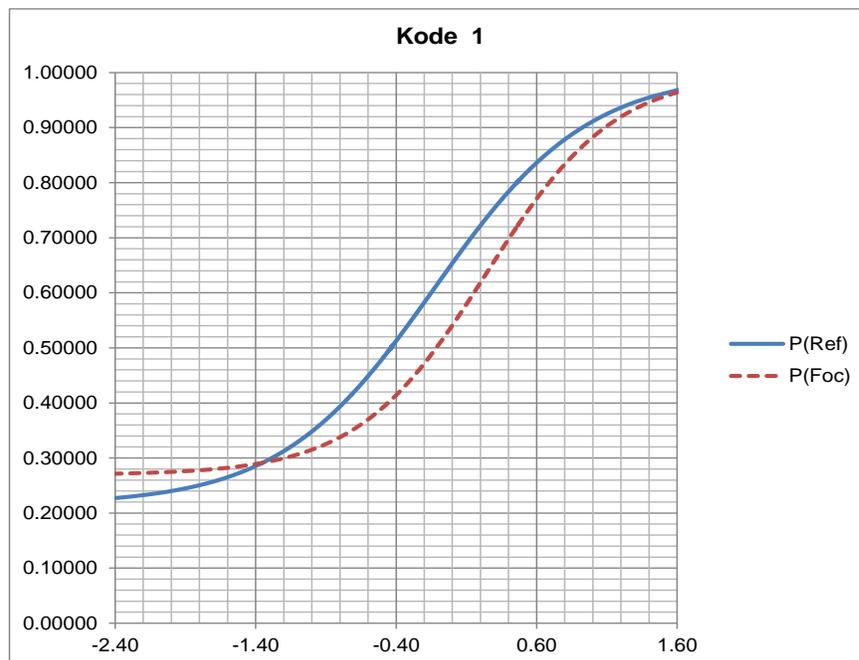
bentuk tersebut merupakan bentuk modus tolens.

Dengan demikian, kesimpulan yang sah dari premis-premis tersebut adalah: Ayah tidak ke rumah sakit.

∴ Kunci Jawaban: B

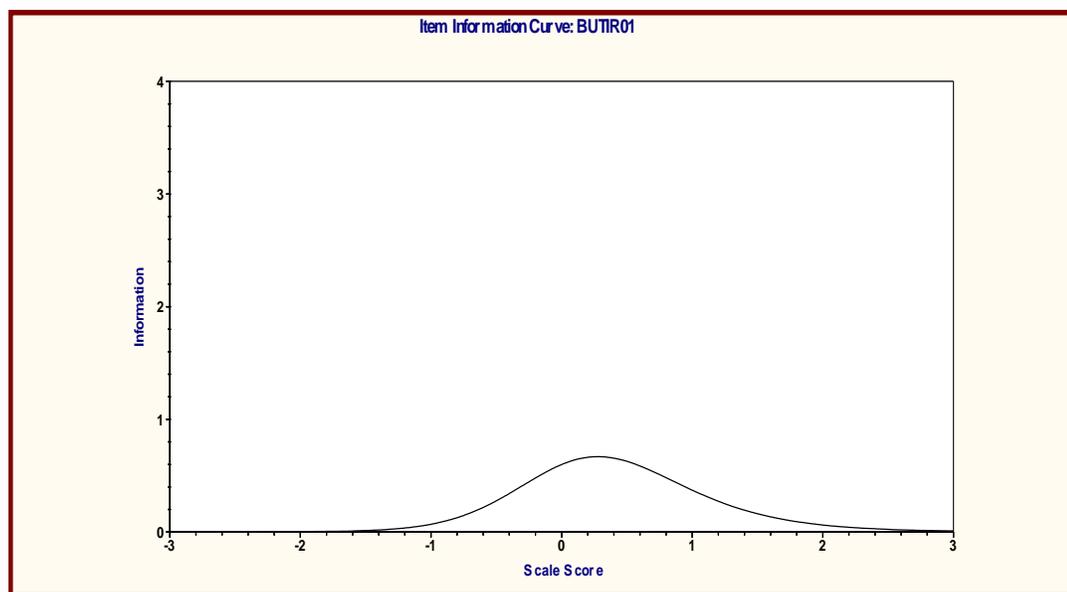
Untuk dapat menjawab butir soal nomor kode 1 tersebut, siswa harus memahami bentuk-bentuk ekuivalen pernyataan majemuk dalam logika matematika. Ada kemungkinan kurang optimalnya kemampuan siswa wanita dalam menjawab butir ini dikarenakan wanita lebih dikuasai emosi sehingga kecemasan muncul dan akibatnya kemampuan menjawab butir ini kurang optimal. Hal ini sesuai dengan pendapat Ulric Neisser et. al., (1996), Setia Iriyanto dan Eny Winaryati (2010) yang menyatakan bahwa perbedaan kemampuan dasar seseorang (intelegenssi) juga dipengaruhi oleh faktor-faktor sosial dan biologis. Perbedaan jenis kelamin dan gender dapat mempengaruhi kemampuan seseorang untuk memecahkan suatu masalah. Peserta didik laki-laki dan perempuan memiliki sejumlah perbedaan yang disebabkan oleh adanya perbedaan perlakuan, kebiasaan, dan perasaan

(emosi). Yeni Tri Asmaningtyas (2012) menyatakan bahwa berkaitan dengan pembelajaran matematika di sekolah yang melibatkan siswa laki-laki dan perempuan, pada umumnya siswa perempuan lebih baik dalam ingatan dan siswa laki-laki lebih baik dalam hal berpikir logis. Oleh karena butir soal nomor 1 mengukur tentang kemampuan logika, maka wajarlah siswa laki-laki (kelompok referensi) kemampuannya lebih baik dibandingkan dengan siswa perempuan (kelompok fokal). Namun pada interval kemampuan $-2,41897 \leq \theta \leq -1,35301$ (kemampuan rendah) butir soal tersebut memberikan keuntungan bagi kelompok fokal. Ada kemungkinan probabilitas menjawab benar tersebut dipengaruhi oleh faktor kebetulan menjawab betul.



Gambar 4.3 Kurva Karakteristik Butir Kelompok Referensi dan Fokal Kode 1

Pada Gambar 4.3 dapat dilihat bahwa butir dengan nomor kode 1 merupakan bias butir jenis *non-uniform*. Pada interval $-2,41897 \leq \theta \leq -1,35301$ kurva karakteristik butir kelompok fokal berada di atas kurva karakteristik butir kelompok referensi. Artinya pada interval tersebut butir soal nomor kode 1 menguntungkan kelompok fokal (perempuan). Sedangkan pada interval $-1,339871 \leq \theta \leq 1,68553$ kurva karakteristik butir kelompok referensi berada di atas kurva karakteristik butir kelompok fokal. Artinya pada interval tersebut butir soal nomor kode 1 menguntungkan kelompok referensi.



Gambar 4.4 Grafik Fungsi Informasi Butir Kode 1

Pada Gambar 4.4 butir kode 1 memiliki nilai fungsi informasi butir maksimum sebesar 0,6670. Nilai fungsi informasi butir sangat dipengaruhi oleh kecuraman grafik dan besarnya penyebaran (simpangan baku).

Kecuraman grafik ditandai oleh nilai parameter daya pembeda (a). Pada butir nomor 1 nilai estimasi parameter daya pembeda untuk kelompok referensi sebesar 1,07353 dan pada kelompok fokal sebesar 1,41973. Termasuk butir soal dengan daya pembeda yang cukup baik, karena nilainya lebih besar dari 1.

Selanjutnya perhatikan soal dengan nomor kode 13 atau butir soal nomor 19, pada replikasi ke-1. Butir soal tersebut juga terdeteksi sebagai bias butir menggunakan ketiga metode *DIF*. Pada kelompok referensi, butir tersebut memiliki parameter $a=2,44150$; $b=0,18874$; $c=0,13830$; dan *standard error* parameter b adalah $S_R=0,042$. Pada kelompok fokal sebelum disetarakan, butir tersebut memiliki parameter $a=2,36275$; $b=0,36723$; $c=0,23473$; dan *standard error* parameter b $S_F=0,050$. Setelah disetarakan menggunakan metode rerata dan sigma, butir tersebut memiliki parameter $a=2,13404$; $b=0,38794$; dan $c=0,23473$. Besarnya konstanta penyetaraan $A=1,10717$; $K=-0,018644$; $\mu_{bR}=-0,20029$; $\mu_{bF}=-0,16406$; $\sigma_{bR}=1,11269$; dan $\sigma_{bF}=1,00499$. Bila dilihat dari parameter daya pembeda, baik untuk kelompok referensi maupun fokal, butir soal nomor 19 memiliki daya pembeda yang sangat baik $a \geq 1$ bahkan nilainya di atas 2. Ditinjau dari parameter taraf sukar, butir soal nomor 19 termasuk soal kategori sedang dengan taraf sukar $-0,99 \leq b \leq 1$. Namun soal ini memiliki nilai kebetulan menjawab betul tidak baik karena nilai c sebesar $0,23473 > 0,200$ untuk kelompok fokal.

Butir soal nomor kode 13 (soal nomor 19) tersebut merupakan soal tentang transformasi geometri. Transformasi yang digunakan dalam soal tersebut merupakan komposisi transformasi. Berikut disajikan indikator, rumusan butir soal, dan pembahasannya.

Indikator : Menentukan bayangan titik atau kurva karena dua transformasi atau lebih.

Soal : Diketahui T_1 adalah transformasi pencerminan terhadap garis $y=x$, dan T_2 adalah rotasi dengan pusat $O(0,0)$ sebesar 90° dengan arah putar berlawanan dengan putaran jarum jam. Persamaan bayangan garis $2x-5y+3=0$ oleh transformasi T_1 dilanjutkan T_2 adalah....

A. $2x+5y-3=0$

B. $2x-5y-3=0$

C. $2x+5y+3=0$

D. $5x-2y-3=0$

E. $5x-2y+3=0$

Pembahasan : Matriks transformasi T_1 , $T_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, dan

$$R_{(0,0)90^\circ} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \Leftrightarrow T_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot T_1$$

dilanjutkan T_2 , sehingga persamaan matriks transformasi adalah:

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \right] = \begin{pmatrix} x' \\ y' \end{pmatrix};$$

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} y \\ x \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix};$$

$$\begin{pmatrix} -x \\ y \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix}$$

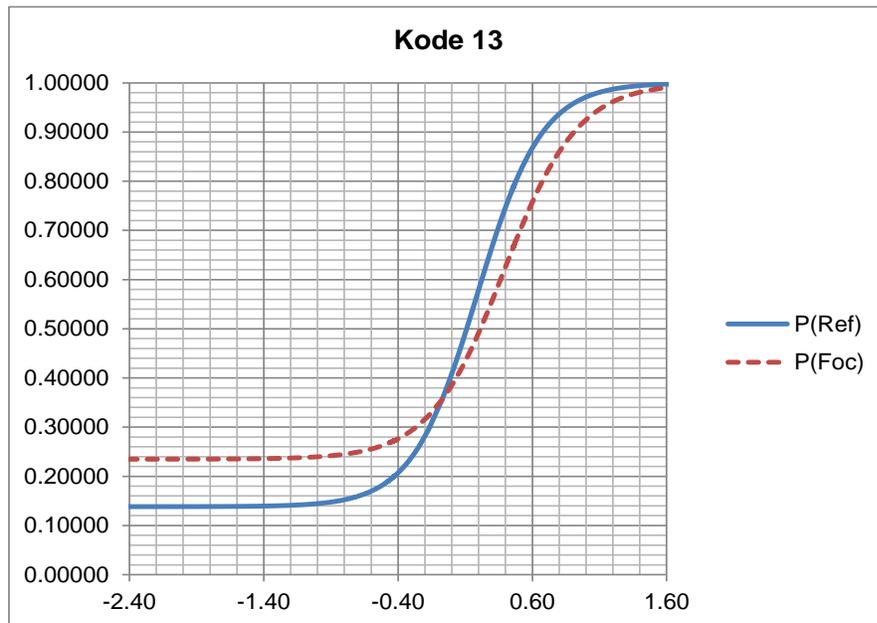
Sehingga bayangan garis $2x-5y+3=0$ adalah $2x'-5y'+3=0$,

dengan: $2(-x)-5(y)+3=0 \Leftrightarrow 2x+5y-3=0$.

∴ Kunci Jawaban: A

Sesuai dengan pendapat Benbov dan Stanley (1992) yang menyatakan bahwa dilihat hubungan jenis kelamin dan hasil belajar matematika, siswa laki-laki lebih unggul dalam matematika terutama dalam soal-soal spasial. Sehingga dalam topik-topik tertentu dalam geometri dan ilmu ukur ruang, hasil belajar siswa laki-laki lebih baik daripada siswa perempuan. Terlihat bahwa pada interval kemampuan $-0,06548 \leq \theta \leq 1,68553$ siswa laki-laki lebih unggul dalam topik geometri dibandingkan dengan siswa perempuan (kelompok fokal). Karena untuk dapat mengerjakan soal ini dengan baik, siswa harus menguasai konsep-konsep geometri spasial terutama tentang garis. Representasi pendukung lainnya adalah pemahaman tentang konsep perkalian matriks, dan sistem persamaan. Pada interval kemampuan rendah, terlihat kelompok fokal (siswa perempuan) lebih unggul bila dibandingkan dengan siswa laki-laki (referensi). Pada siswa dengan kemampuan rendah, pada umumnya mereka pasrah akibat ketidakmampuan menjawab soal menggunakan konsep geometri. Dengan demikian ada kemungkinan, siswa dapat menjawab dengan benar karena faktor kebetulan menjawab betul (*pseudoguessing*). Hal ini dapat dilihat dari faktor kebetulan menjawab betul tidak baik karena nilai c sebesar

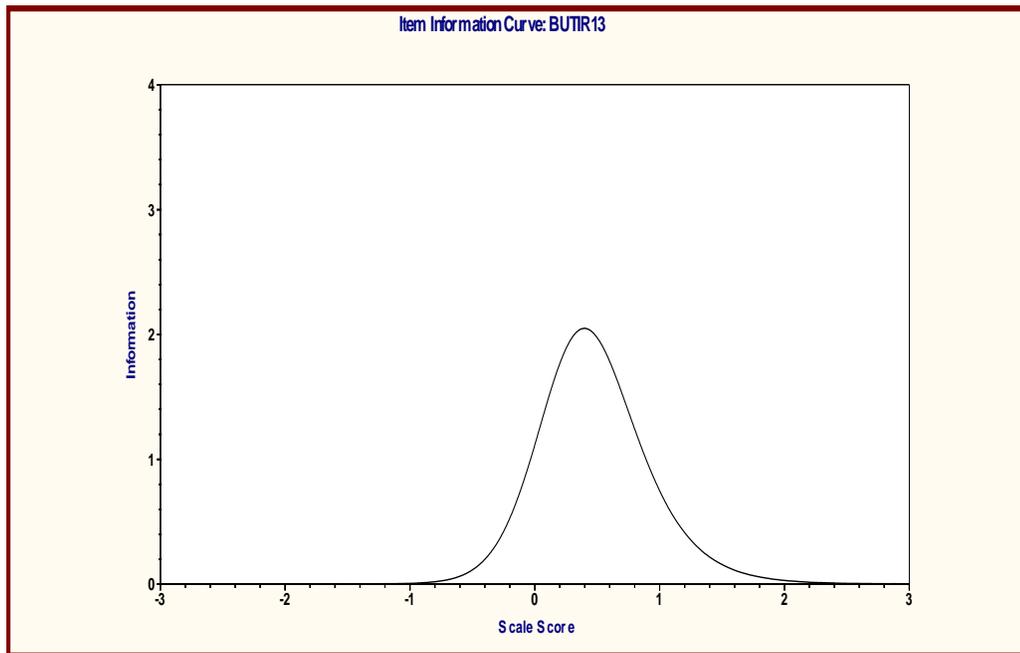
0,23473 > 0,200 (untuk 5 pilihan jawaban) untuk kelompok fokal (siswa perempuan).



Gambar 4.5 kurva Karakteristik Butir
Kelompok Referensi dan Fokal Kode 13

Pada Gambar 4.5 dapat dilihat bahwa butir dengan nomor kode 13 merupakan bias butir jenis *non-uniform*. Pada interval $-2.41897 \leq \theta \leq -0.07512$ kurva karakteristik butir kelompok fokal berada di atas kurva karakteristik butir kelompok referensi. Artinya pada interval tersebut butir soal nomor kode 13 menguntungkan kelompok fokal (perempuan). Sedangkan pada interval $-0,06548 \leq \theta \leq 1,68553$ kurva karakteristik butir kelompok referensi berada di atas kurva karakteristik butir kelompok fokal. Artinya pada

interval tersebut butir soal nomor kode 13 menguntungkan kelompok referensi.



Gambar 4.6 Grafik Fungsi Informasi Butir Kode 13

Pada Gambar 4.6 butir kode 13 memiliki nilai fungsi informasi butir maksimum sebesar 2,0469. Nilai fungsi informasi butir sangat dipengaruhi oleh kecuraman grafik dan besarnya penyebaran (simpangan baku). Kecuraman grafik ditandai oleh nilai parameter daya pembeda (a). Pada butir nomor 13 nilai estimasi parameter daya pembeda untuk kelompok referensi sebesar 2,44150; dan pada kelompok fokal sebesar 2,36275. Termasuk butir soal dengan daya pembeda yang sangat baik, karena nilainya lebih besar dari 2.

Berikutnya dibahas soal dengan nomor kode 24 atau butir soal nomor 34, pada replikasi ke-1. Butir soal tersebut juga terdeteksi sebagai bias butir menggunakan ketiga metode *DIF*. Pada kelompok referensi, butir tersebut memiliki parameter $a=1,74666$; $b=-0,19992$; $c=0,12348$; dan *standard error* parameter b adalah $S_R=0,060$. Pada kelompok fokal sebelum disetarakan, butir tersebut memiliki parameter $a=2,12401$; $b=0,01097$; $c=0,15966$; dan *standard error* parameter b $S_F=0,053$. Setelah disetarakan menggunakan metode rerata dan sigma, butir tersebut memiliki parameter $a=1,91841$; $b=-0,00650$; dan $c=0,15966$. Besarnya konstanta penyetaraan $A=1,10717$; $K=-0,018644$; $\mu_{bR}=-0,20029$; $\mu_{bF}=-0,16406$; $\sigma_{bR}=1,11269$; dan $\sigma_{bF}=1,00499$. Bila dilihat dari parameter daya pembeda, baik untuk kelompok referensi maupun fokal, butir soal nomor 34 memiliki daya pembeda yang sangat baik $a \geq 1$. Ditinjau dari parameter taraf sukar, butir soal nomor 34 termasuk soal kategori sedang dengan taraf sukar $-0,99 \leq b \leq 1$. Demikian pula soal ini memiliki nilai kebetulan menjawab betul yang baik karena nilai c baik untuk kelompok referensi dan fokal lebih kecil daripada 0,200.

Soal nomor kode 24 tersebut merupakan soal tentang integral trigonometri. Berikut disajikan rincian indikator, rumusan butir soal dan pembahasannya.

Indikator : Menentukan integral tak tentu dan integral tentu fungsi aljabar dan fungsi trigonometri.

Soal : Hasil $\int 4 \sin 4x \cos 2x \, dx$ adalah....

A. $-\frac{1}{6} \cos 6x - \frac{1}{2} \cos 2x + C$

B. $-\frac{1}{3} \cos 6x - \cos 2x + C$

C. $\frac{1}{6} \cos 6x - \frac{1}{2} \cos 2x + C$

D. $\frac{1}{6} \cos 6x + \frac{1}{2} \cos 2x + C$

E. $-\frac{1}{6} \cos 6x + \cos 2x + C$

Pembahasan : Bentuk: $2 \sin A \cos B = \sin (A+B) + \sin (A-B)$, sehingga bentuk di atas menjadi:

$$2 \int 2 \sin 4x \cos 2x \, dx = 2 \int (\sin 6x + \sin 2x) \, dx$$

$$2 \int \sin 6x \, dx + 2 \int \sin 2x \, dx = 2 \cdot \left(-\frac{1}{6}\right) \cos 6x$$

$$+ 2 \cdot \left(-\frac{1}{2}\right) \cos 2x + C$$

$$= (-1/3) \cos 6x - \cos 2x + C$$

∴ Kunci Jawaban: B

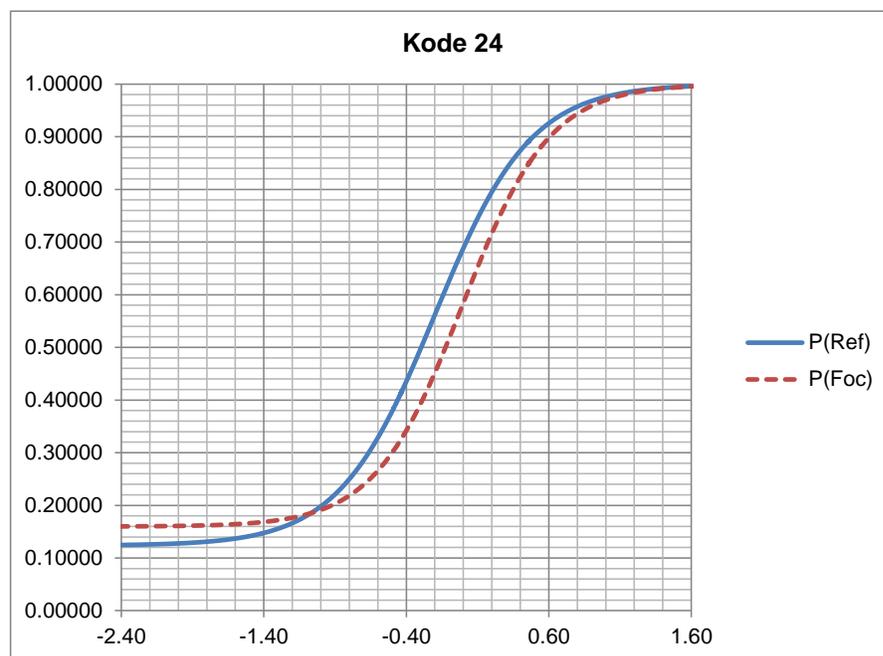
Butir soal nomor 34 (kode 24) menguji kemampuan menyelesaikan integral tak tentu dan integral tentu fungsi aljabar dan fungsi trigonometri. Dari topik yang diujikan, terlihat soal tersebut memiliki karakteristik sangat teoretis dan abstrak. Agar dapat menyelesaikan soal tersebut dengan baik, diperlukan banyak kompetensi meliputi logika yang tinggi untuk mengubah bentuk integral perkalian menjadi integral penjumlahan. Termasuk

kemampuan mengingat rumus-rumus dasar identitas trigonometri dan rumus anti diferensial (integral).

Menurut pendapat Kartini Kartono (1989) dalam Yeni Tri Asmaningtyas (2012), betapapun baik dan cemerlangnya intelegensi perempuan, namun pada intinya perempuan hampir-hampir tidak pernah mempunyai ketertarikan yang menyeluruh pada soal-soal teoretis seperti pada laki-laki. Siswa perempuan lebih tertarik pada hal-hal yang praktis dan konkret. Sedangkan siswa laki-laki lebih tertarik pada soal-soal yang teoretis dan abstrak. Hasil penelitian ini menunjukkan bahwa pada level kemampuan yang tinggi kemampuan siswa laki-laki untuk menyelesaikan soal-soal yang teoretik dan abstrak seperti integral lebih baik bila dibandingkan dengan kemampuan siswa perempuan. Pada interval $-1,05835 \leq \theta \leq 1,68553$ kelompok siswa laki-laki (referensi) lebih unggul prestasinya menyelesaikan soal integral trigonometri bila dibandingkan dengan kelompok siswa perempuan (fokal). Sedangkan pada kelompok siswa dengan kemampuan rendah, pada umumnya mereka menggunakan naluri (menebak jawaban) untuk menjawab soal. Dengan demikian, apabila tebakan mereka tepat, mereka memiliki peluang menjawab benar yang sama dengan kelompok lainnya. Terlihat bahwa nilai faktor kebetulan menjawab betul (c) untuk kedua kelompok referensi dan fokal berturut-turut $c=0,12348$ dan $c=0,15966$. Terlihat bahwa faktor kebetulan menjawab betul untuk kelompok fokal lebih besar daripada

kelompok referensi. Hal itu berarti bahwa, lebih tingginya prestasi siswa perempuan (kelompok focal) daripada siswa laki-laki (kelompok referensi) pada level kemampuan rendah disebabkan oleh kemampuan menebak.

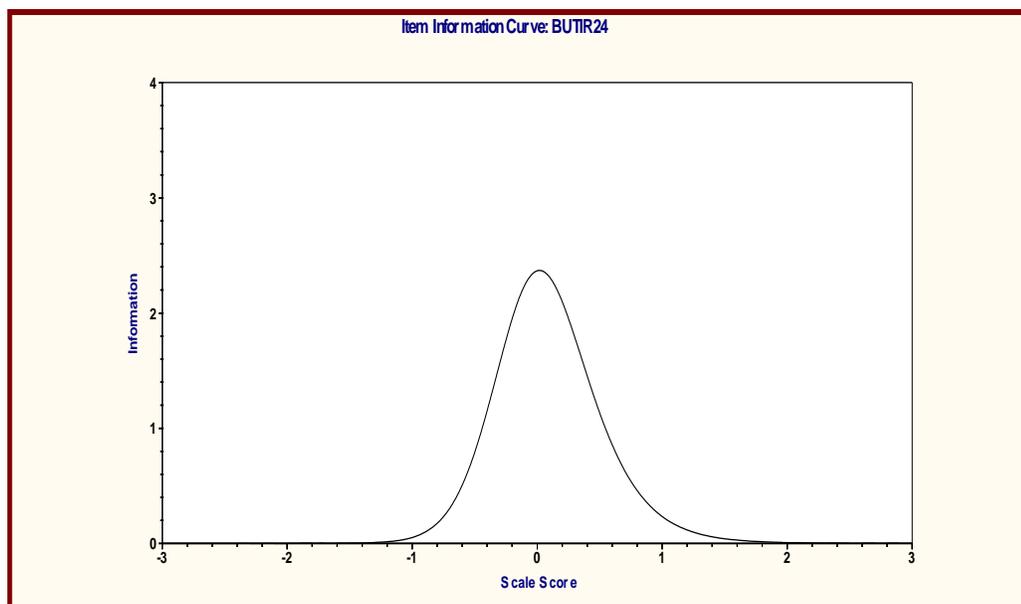
Perbandingan kurva karakteristik butir untuk kelompok referensi dan focal butir soal dengan nomor kode 24, disajikan pada gambar 4.7 berikut.



Gambar 4.7 Kurva Karakteristik Butir
Kelompok Referensi dan Fokal Kode 24

Pada Gambar 4.7 terlihat bahwa butir dengan nomor kode 24 merupakan bias butir jenis *non-uniform*. Pada interval $-2,41897 \leq \theta \leq -1,07145$ kurva karakteristik butir kelompok focal berada di atas kurva karakteristik butir kelompok referensi. Artinya pada interval tersebut butir soal

nomor kode 24 menguntungkan kelompok fokal (perempuan). Sedangkan pada interval $-1,05835 \leq \theta \leq 1,68553$ kurva karakteristik butir kelompok referensi berada di atas kurva karakteristik butir kelompok fokal. Artinya pada interval tersebut butir soal nomor kode 24 menguntungkan kelompok referensi (laki-laki).



Gambar 4.8 Grafik Fungsi Informasi Butir Kode 24

Butir soal dengan nomor kode 24 memiliki nilai fungsi informasi butir maksimum sebesar 2,3692. Pada butir nomor 24 nilai estimasi parameter daya pembeda untuk kelompok referensi sebesar 1,74666; dan pada kelompok fokal sebesar 2,12401. Termasuk butir soal dengan daya pembeda yang sangat baik.

Walaupun hasil penelitian ini telah membuktikan bahwa metode DSE lebih sensitif untuk mendeteksi bias butir bila dibandingkan dengan metode Khi-Kuadrat Lord, dan metode Khi-Kuadrat Lord lebih sensitif mendeteksi bias butir bila dibandingkan dengan metode UBTS. Namun, dalam penelitian ini juga diperoleh hasil bahwa, terdapat beberapa butir yang dinyatakan bias menggunakan metode UBTS, tetapi dinyatakan tidak bias oleh metode Khi-Kuadrat Lord. Demikian pula butir yang dinyatakan bias menggunakan metode Khi-Kuadrat Lord, tetapi dinyatakan tidak bias menggunakan metode DSE. Hasil-hasil tersebut dibahas dalam pembahasan berikut.

Masing-masing metode *DIF* yang diperbandingkan memiliki keunggulan dan kelemahan. Sebagaimana telah diuraikan sebelumnya bahwa metode UBTS dalam mendeteksi bias butir menggunakan rumus:

$$d = \frac{\Delta \hat{b}}{\sqrt{S_R^2 + S_F^2}}$$

Apabila nilai $\Delta \hat{b}$ cukup besar dan $\sqrt{S_R^2 + S_F^2}$ cukup kecil, maka nilai

d akan cukup besar. Sedangkan jika nilai $\Delta \hat{b}$ diasumsikan konstan, maka nilai d dipengaruhi oleh nilai $\sqrt{S_R^2 + S_F^2}$. Semakin besar nilai $\sqrt{S_R^2 + S_F^2}$ maka nilai d semakin kecil, dan sebaliknya semakin kecil nilai $\sqrt{S_R^2 + S_F^2}$, maka nilai d semakin besar. Hal itu berarti bahwa semakin kecil nilai *standard error* estimasi parameter b , baik untuk kelompok referensi dan kelompok fokal maka nilai d semakin besar. Artinya semakin besar kemungkinan butir tersebut dinyatakan bias oleh metode UBTS. Demikian pula bila $\sqrt{S_R^2 + S_F^2}$

diasumsikan konstan, maka besarnya nilai d dipengaruhi oleh $\Delta\hat{b}$. Semakin besar nilai $\Delta\hat{b}$, semakin besar nilai d dan sebaliknya semakin kecil nilai $\Delta\hat{b}$ maka semakin kecil pula nilai d . Hal itu berarti bahwa semakin besar perbedaan estimasi parameter b pada kelompok referensi dan fokal, mengakibatkan nilai d semakin besar sehingga semakin besar kemungkinan butir tersebut dinyatakan bias oleh metode UBTS.

Di sisi lain metode UBTS memiliki kelemahan yaitu tidak mempertimbangkan perbedaan parameter a dan c antara kelompok referensi dan fokal. Sehingga apabila terjadi perbedaan nilai yang cukup besar pada parameter a dan c , tidak mempengaruhi hasil analisis bias butir menurut metode UBTS. Padahal perbedaan nilai-nilai a dan c pada kelompok referensi dan fokal itu mengakibatkan perbedaan kurva karakteristik butir, yang memberikan kemungkinan butir itu bisa bias. Itulah sebabnya, dalam beberapa replikasi, butir-butir yang seharusnya bias menurut metode lain dinyatakan tidak bias oleh metode UBTS.

Berbeda dengan metode Khi-Kuadrat Lord, dalam perhitungan analisis bias butir menggunakan rumus: $\chi_i^2 = v_i' \Sigma_i^{-1} v_i$. Nilai χ^2 dipengaruhi oleh perbedaan estimasi parameter butir dan invers matriks varians kovarians. Metode Khi-Kuadrat Lord tidak mempertimbangan *standard error* ($\sqrt{S_R^2 + S_F^2}$) pada estimasi b sebagaimana yang dilakukan pada metode UBTS. Artinya pada butir-butir soal dengan *standard error* yang sangat kecil, pada kelompok

referensi dan fokal, menurut metode UBTS dinyatakan sebagai bias butir asalkan nilai $\Delta\hat{b}$ cukup besar. Kemungkinan oleh metode Khi-Kuadrat Lord tidak terdeteksi apabila elemen-elemen pada matriks varians kovarians butir tersebut cukup besar, yang mengakibatkan nilai χ^2 cukup kecil, bahkan lebih kecil daripada nilai χ^2 kritis.

Keunggulan metode UBTS tersebut di atas juga dapat mengakibatkan butir-butir yang dinyatakan bias oleh metode UBTS bisa saja tidak terdeteksi oleh metode DSE, walaupun dalam perhitungan analisis butir melibatkan replikasi analisis. Pada metode DSE, bias butir didefinisikan sebagai perbedaan peluang menjawab benar responden dengan kemampuan yang sama, namun berasal dari kelompok yang berbeda. Indeks bias butir dihitung

menggunakan rumus: $SPD - \theta = \frac{\sum_{j=1}^{n_p} \Delta P(\theta_j)}{n_p}$ untuk bias butir jenis *uniform*

(konsisten) dan $UPD - \theta = \sqrt{\frac{\sum_{j=1}^{n_p} (\Delta P(\theta_j))^2}{n_p}}$ untuk bias butir jenis *non-uniform*

(tidak konsisten). Seperti halnya pada metode Khi-Kuadrat Lord, apabila dalam sebuah butir nilai *standard error* ($\sqrt{S_R^2 + S_F^2}$) pada estimasi b sangat kecil pada kelompok referensi dan fokal, menurut metode UBTS dinyatakan sebagai bias butir asalkan nilai $\Delta\hat{b}$ cukup besar. Kemungkinan oleh metode DSE tidak terdeteksi, karena teknik perhitungan analisis bias butir yang berbeda.

Demikian pula butir-butir yang terdeteksi bias oleh metode Khi-Kuadrat Lord, ada kemungkinan tidak terdeteksi oleh metode DSE karena komponen yang dianalisis berbeda. Pada metode Khi-Kuadrat Lord, komponen perbedaan estimasi parameter butir (a, b, dan c) serta matriks varians kovarian yang mempengaruhi nilai χ^2 . Sedangkan pada metode DSE bias butir ditentukan oleh perbedaan probabilitas menjawab benar responden dengan kemampuan yang sama, namun berasal dari kelompok yang berbeda. Untuk menghitung probabilitas menjawab benar ditentukan oleh estimasi parameter butir dan kemampuan, serta melalui replikasi analisis.

Untuk memperjelas uraian di atas berikut ini dibahas sebuah butir yang terdeteksi bias oleh metode UBTS, tetapi tidak terdeteksi bias oleh metode Khi-Kuadrat Lord dan metode DSE. Pada replikasi ke-18, butir soal dengan nomor kode 23 dinyatakan bias oleh metode UBTS, tetapi tidak terdeteksi bias oleh metode Khi-Kuadrat Lord dan DSE. Karakteristik butir soal kode 23 adalah sebagai berikut: (a) parameter butir untuk kelompok referensi $b = -0,25657$; dan $S_R = 0,066$; (b) parameter butir untuk kelompok fokal (setelah disetarakan) $b = -0,07559$; dan $S_F = 0,061$. Dari data di atas diperoleh $\Delta \hat{b} = -0,18098$; nilai $\sqrt{S_R^2 + S_F^2} = 0,08977$ (sangat kecil di bawah 0,09), dan nilai $d = -2,01618$ lebih kecil daripada nilai kritis $z = -1,96$. Kesimpulannya, butir soal dengan kode 23 dinyatakan bias butir oleh metode UBTS.

Selanjutnya butir soal kode 23 dianalisis menggunakan metode Khi-Kuadrat Lord. Karakteristik butir soal kode 23 adalah sebagai berikut: (a) parameter butir untuk kelompok referensi $a=2,02215$; $b=-0,25657$; $c=0,19570$;

matriks varians kovarians $\Sigma = \begin{bmatrix} 0,05255 & 0,00939 & 0,00465 \\ 0,00939 & 0,00423 & 0,00239 \\ 0,00465 & 0,00239 & 0,00140 \end{bmatrix}$ (b) parameter

butir untuk kelompok fokal (setelah disetarakan) $a=2,40662$; $b=-0,07559$;

$c=0,25271$; matriks varians kovarians $\Sigma = \begin{bmatrix} 0,10790 & 0,01222 & 0,00587 \\ 0,01222 & 0,00367 & 0,00123 \\ 0,00587 & 0,00123 & 0,00104 \end{bmatrix}$.

Dari data tersebut diperoleh: vektor matriks perbedaan parameter butir

$v = \begin{bmatrix} 0,38447 \\ 0,18098 \\ 0,05701 \end{bmatrix}$; $v' = [0,38447 \quad 0,18098 \quad 0,05701]$; dan $\chi^2 = 5,07212$ lebih

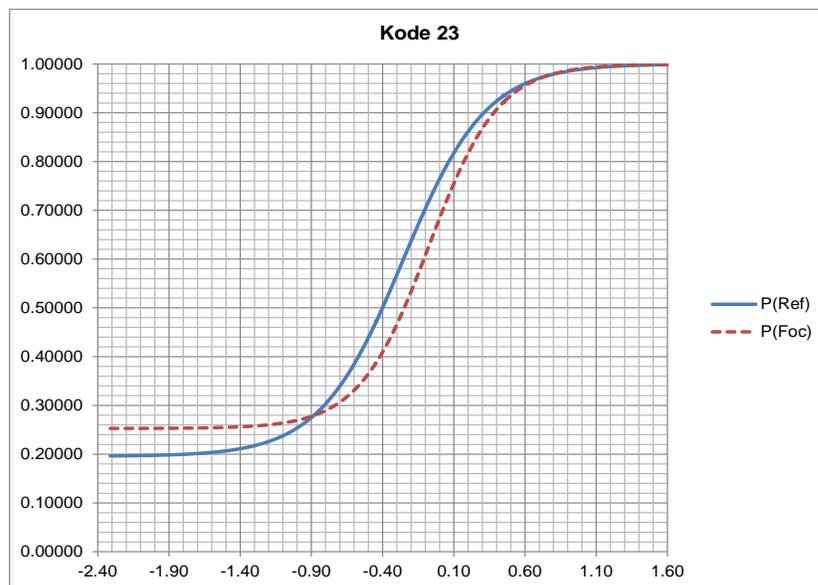
kecil dari nilai χ^2 kritis= $7,81472$. Kesimpulan, butir soal kode 23 tidak bias.

Analisis bias butir soal nomor kode 23 menggunakan metode DSE, dilakukan sebagai berikut. Karakteristik butir soal kode 23 adalah sebagai berikut: (a) parameter butir untuk kelompok referensi $a=2,02215$; $b=-0,25657$; $c=0,19570$; (b) parameter butir untuk kelompok fokal (setelah disetarakan) $a=2,40662$; $b=-0,07559$; $c=0,25271$; (c) parameter butir untuk subkelompok referensi, R1: $a=1,91166$; $b=-0,15119$; $c=0,21504$; R2 (setelah disetarakan): $a=2,38067$; $b=-0,33610$; $c=0,16599$; (d) parameter butir untuk subkelompok fokal, F1: $a=1,95696$; $b=-0,02490$; $c=0,23995$; F2 (setelah disetarakan): $a=2,17391$; $b=0,07330$; $c=0,23005$. Selanjutnya dihitung indeks bias butir

menggunakan rumus: $SPD - \theta = \frac{\sum_{j=1}^{n_p} \Delta P(\theta_j)}{n_p}$ untuk bias butir jenis *uniform*

(konsisten) dan $UPD - \theta = \sqrt{\frac{\sum_{j=1}^{n_p} (\Delta P(\theta_j))^2}{n_p}}$ untuk bias butir jenis *non-uniform*

(tidak konsisten). Hasil perhitungan yang diperoleh adalah: $\delta=0,04841$; $\delta_R=0,05146$; $\delta_F=0,03586$. Oleh karena nilai δ lebih kecil daripada nilai maksimum (δ_R, δ_F), maka kesimpulannya butir soal kode 23 dinyatakan tidak bias oleh metode DSE. Artinya indeks perbedaan probabilitas menjawab benar untuk responden dari kelompok referensi dan fokal tidak signifikan. Berikut ini disajikan perbandingan kurva karakteristik butir kelompok referensi dan fokal untuk butir kode 23 pada replikasi ke-18.



Gambar 4.9 Kurva Karakteristik Butir
Kelompok Referensi dan Fokal Kode 23

Dengan demikian jelaslah bahwa adanya beberapa butir yang dinyatakan bias oleh metode UBTS tetapi tidak terdeteksi bias oleh metode Khi-Kuadrat Lord dan metode DSE, atau butir yang dinyatakan bias oleh metode Khi-Kuadrat Lord tidak terdeteksi bias oleh metode DSE karena adanya perbedaan teknik dalam perhitungan indeks bias butir. Kondisi tersebut terjadi sangat spesifik pada butir-butir tertentu saja. Secara umum, butir-butir yang dinyatakan bias oleh metode UBTS, juga terdeteksi bias oleh metode Khi-Kuadrat Lord dan metode DSE.

Penelitian ini memiliki beberapa keterbatasan yang terkait dengan replikasi, jumlah sampel, dan program aplikasi yang digunakan untuk menganalisis data. Keterbatasan-keterbatasan tersebut hendaknya menjadi perhatian bagi para peneliti dan pengguna hasil penelitian ini dalam pengembangan penelitian sejenis di masa yang akan datang. Beberapa keterbatasan dalam penelitian ini adalah sebagai berikut.

Pertama, pada uji kecocokan model tidak menggunakan replikasi terhadap butir-butir yang cocok dengan model. Apabila dilakukan replikasi, ada kemungkinan perbedaan butir-butir yang cocok dengan model, karena nilai **CHISQ** dan probabilitasnya pada **output.PH2** Program BILOG-MG dapat berubah sesuai dengan respons peserta tes yang dijadikan sampel dalam replikasi.

Kedua, jumlah sampel respons pada uji kecocokan model dalam penelitian ini sebanyak 700 orang responden. Jumlah tersebut dianggap sudah cukup untuk model L3P, didasarkan atas hasil uji nilai *Kaiser-Meyer-Olkin Measure of Sampling Adequacy* untuk model L3P, sebesar 0,940 lebih besar dari 0,5 dengan $\text{sig.} 0,000 < 0,05$ (signifikan) yang berarti bahwa jumlah sampel telah cukup. Jumlah sampel dapat mempengaruhi jumlah butir yang cocok dengan model L3P. Ada kecenderungan bahwa semakin banyak jumlah sampel respons yang digunakan, semakin banyak butir yang tidak cocok dengan model L3P. Hal ini sesuai dengan pendapat Dali S. Naga (1992) yang menyatakan bahwa pada ukuran sampel peserta yang kecil semua butir cenderung diterima, sedangkan pada ukuran sampel peserta yang besar banyak butir yang cenderung ditolak.⁹³

Ketiga, perhitungan analisis bias butir pada metode UBTS, Khi-Kuadrat Lord, dan DSE dilakukan menggunakan Program Excel 2010. Sedangkan estimasi parameter butir dan kemampuan serta varians kovarians kelompok referensi dan fokal dianalisis menggunakan Program BILOG-MG, hasil estimasi tersebut dinyatakan dalam format Notepad. Agar hasil estimasi parameter tersebut dapat diolah menggunakan Program Excel 2010, terlebih dahulu hasil analisis Program BILOG-MG dalam bentuk Notepad diubah dalam bentuk format Excel. Cara ini terkesan kurang efisien terutama pada

⁹³ Dali S. Naga, *op. cit*, h. 305.

metode DSE karena analisis dilakukan secara bertahap (replikasi analisis), sehingga memerlukan banyak langkah serta diikuti dengan perubahan format dari Notepad ke Excel dan sebaliknya harus dilakukan berkali-kali.

Keempat, pada setiap replikasi pemilihan sampel respons dilakukan secara acak menggunakan Program Excel 2010. Pada beberapa pengacakan sampel, Program BILOG-MG mengalami *trouble* saat mengestimasi parameter (*Run Incomplit*), sehingga proses estimasi parameter tidak dapat diselesaikan. Ada kemungkinan proses pengacakan yang dilakukan oleh Program Excel tidak dapat diterima oleh Program BILOG-MG, karena memiliki pola yang sama dengan pengacakan sebelumnya, sehingga harus diulangi. Belum diketahui secara pasti dalam penelitian ini, berapa kali pengacakan pemilihan sampel dengan jumlah tertentu dapat dilakukan, pada populasi dengan jumlah tertentu pula, sehingga menghasilkan pola respons berbeda yang dapat dianalisis oleh Program BILOG. Kemungkinan kedua, disebabkan karena pada saat perubahan format data dari Excel ke Notepad terjadi *trouble* dalam *sytem* sehingga BILOG-MG tidak dapat mengolah data yang diubah dari format Excel ke Notepad dengan baik.

Kelima, perhitungan analisis bias butir pada metode UBTS, Khi-Kuadrat Lord, dan DSE dalam penelitian ini menggunakan jumlah sampel 1000 orang untuk masing-masing kelompok referensi dan fokal. Khusus pada

metode DSE, analisis bias butir dilakukan secara berjenjang (replikasi analisis) dengan teknik membagi dua masing-masing kelompok referensi dan fokal. Dengan demikian jumlah sampel pada masing-masing subkelompok adalah 500 orang. Ada kemungkinan apabila jumlah sampel pada kelompok referensi dan fokal diperbesar (lebih dari 1000 orang) atau diperkecil (kurang dari 1000 orang) mengakibatkan adanya perbedaan nilai estimasi parameter butir. Perbedaan hasil estimasi parameter butir tersebut dapat mengakibatkan perbedaan hasil analisis bias butir pada masing-masing metode *DIF*. Sehingga perlu diteliti lebih lanjut, untuk mendapatkan jumlah sampel yang representatif agar perbedaan nilai estimasi parameter butir relatif stabil.

BAB V

KESIMPULAN, IMPLIKASI DAN SARAN

Berdasarkan hasil penelitian dan pembahasan, pada bagian terakhir ini akan dipaparkan berturut-turut kesimpulan, implikasi, dan saran. Secara rinci, kesimpulan, implikasi, dan saran penelitian ini akan diuraikan sebagai berikut.

A. Kesimpulan

Penelitian ini membandingkan sensitivitas metode *DIF* untuk mendeteksi bias butir. Metode *DIF* yang diperbandingkan adalah metode Uji Beda Taraf Sukar (UBTS), metode Khi-Kuadrat Lord, dan metode Distribusi Sampling Empiris (DSE). Bias butir adalah butir yang mengakibatkan adanya skor tes yang berbeda bila diujikan pada peserta dengan kemampuan yang sama, namun mereka berasal dari kelompok yang berbeda.

Sensitivitas metode *DIF* adalah kemampuan suatu metode *DIF* untuk mendeteksi adanya bias butir dalam suatu perangkat tes. Sensitivitas metode *DIF* dapat dilihat dari banyaknya jumlah butir yang terdeteksi mengandung bias butir oleh metode *DIF*. Berdasarkan hasil penelitian dapat ditarik beberapa kesimpulan sebagai berikut.

1. Metode Distribusi Sampling Empiris (DSE) lebih sensitif daripada metode Khi-Kuadrat Lord untuk mendeteksi bias butir.

2. Metode Khi-Kuadrat Lord lebih sensitif daripada metode Uji Beda Taraf Sukar (UBTS) untuk mendeteksi bias butir.

B. Implikasi

Secara umum, implikasi hasil penelitian ini dapat dipilah menjadi dua bagian, yaitu implikasi teoretis dan praktis.

1. Implikasi teoretis yang dapat dikembangkan dari hasil penelitian ini adalah: (a) hasil penelitian ini dapat memperkuat teori-teori tentang sensitivitas metode *DIF* yang telah dikembangkan dan diteliti sebelumnya, (b) menambah wawasan dan referensi tentang kajian ilmu pengukuran terkait dengan metode *DIF* dan bias butir.
2. Implikasi praktis yang dapat dikembangkan dari hasil penelitian ini adalah: (a) metode *DIF* yang diperbandingkan sensitivitasnya dalam penelitian ini dapat digunakan sebagai salah satu metode alternatif untuk mendeteksi bias butir pada Ujian Sekolah (US), seleksi Penerimaan Peserta Didik Baru (PPDB), atau penilaian dalam skala yang lebih luas lainnya seperti Ujian Nasional (UN) dan Seleksi Bersama Masuk Perguruan Tinggi Negeri (SBMPTN) untuk meminimalkan adanya kelompok-kelompok tertentu yang dirugikan atau diuntungkan oleh perangkat tes, (b) metode *DIF* yang telah diperbandingkan sensitivitasnya dalam penelitian ini, dapat dijadikan penelitian lanjutan terutama yang berkaitan dengan keterbatasan-keterbatasan penelitian ini.

C. Saran

Sesuai dengan hasil penelitian dan keterbatasan dalam penelitian ini, dapat diajukan beberapa saran sebagai berikut.

1. Bagi *stakeholders* (pemangku kebijakan) seperti Pusat Penilaian Pendidikan (Puspendik), Kepala Dinas Pendidikan, Kepala Sekolah, dan Guru disarankan untuk melakukan uji bias butir terhadap perangkat tes yang akan diujikan sehingga dapat meminimalkan terjadinya bias butir yang dapat menguntungkan atau merugikan pihak-pihak tertentu.
2. Bagi para penulis soal, agar mempertimbangkan serta meminimalkan faktor-faktor yang memungkinkan terjadinya bias butir akibat perbedaan gender, suku, wilayah, agama, etnik atau golongan tertentu.
3. Bagi para peneliti disarankan: (a) melakukan penelitian lanjutan untuk mengembangkan hasil penelitian ini pada mata pelajaran lainnya, (b) mengembangkan *software* yang terkait dengan metode *DIF* UBTS, *Khi-Kuadrat* Lord, dan *DSE* untuk mendeteksi bias butir yang dapat digunakan dengan mudah, praktis, dan presisi. Dalam penelitian ini masih menggunakan beberapa *software*, sehingga terkesan kurang praktis dan langkah yang dilakukan dirasakan berbelit-belit.

DAFTAR PUSTAKA

- Adam E. Wyse & Raymond Mapuranga. Differential Item Functioning Analysis Using Rasch Item Information Functions. *International Journal of Testing*, 9: 333–357, 2009, Taylor & Francis Group, LLC, ISSN: 1530-5058 (diakses 9 Oktober 2015).
- Agung, I Gusti Ngurah. *Statistika Penerapan Model Rerata Sel Multivariat dan Model Ekonometri dengan SPSS*. Jakarta: Yayasan Sad Satria Bhakti, 2006.
- _____. *Manajemen Penulisan Skripsi, Tesis, dan Disertasi*. Jakarta: PT. Raja Grafindo Persada, 2004.
- Allan S. Cohen & Seok-Ho Kim. An Investigation of Linking Method Under The Graded Response Model. *Journal Applied Psychological Measurement*, Vol. 22, No. 2, Juni 1998.
- Allen, M. J. & Yen, W. M. *Introduction to Measurement Theory*. Belmont, CA: Wadsworth Inc, 1979.
- Camilli, G. dan Shepard, L. A., *Methods for Identifying Biased Test Item*. California: Sage Publications Inc., 1994.
- Carlo Magno, Demonstrating the Difference between Classical Test Theory and Item Response Theory Using Derived Test Data. *The International Journal of Educational and Psychological Assessment April 2009, Vol. 1, Issue 1, pp. 1-11* (diakses 5 Maret 2016).
- Darmastini, Dian Prastiwi dan Rosyidi, Abdul Haris. *Multi Representasi Siswa SMP dalam Menyelesaikan Soal Terbuka Matematika Ditinjau dari Perbedaan Gender. Jurnal Ilmiah Pendidikan Matematika MATHEdunesa, Volume 3 Nomor 1 Tahun 2014.* <http://ejournal.unesa.ac.id/article/9928/30/article.doc> (diakses 10 Oktober 2015).
- Deping Li, Yanlin Jiang, dan Alina A. von Davier, The Accuracy and Consistency of a Series of IRT True Score Equatings. *Journal of Educational Measurement Summer 2012, Vol. 49, No. 2, pp. 167–189* (diakses 5 Maret 2016)

- Effendi. Pendeteksian *Crossing DIF*: Suatu Komparasi Metode Luasan Menurut Raju Chi-Square Menurut Lord, dan Tes Rasio Kebolehjadian. *Disertasi*. Jakarta: UNJ, 2011.
- Embretson, S. E., & Reise, S. P. *Item Response Theory for Psychologists*. NJ: Lawrence Erlbaum Associates Inc., 2000.
- Finch, W. Holmes dan Brian F. French. Detection of Crossing Differential Item Functioning. A Comparison of Four Methods. *Journal Education and Psychological Measurement*, volume 67, No. 4, 2007.
- Hambleton, R. K., dan Swaminathan H. *Item Response Theory*. Boston: Kluwer Nijhoff Publishing, 1985.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J., *Fundamentals of Item Response Theory*. CA: Sage Publication Inc., 1991.
- Hog, R. V., & Craig, A. T., *Introduction to Mathematical Statistics 4th ed*. New York: Collier Macmillan, 1978.
- Hulin, C.L., Drasgow, F., & Parsons, C. K., *Item Response Theory: Application to Psychological Measurement*. Homewood, IL: Dow Jones-Irwin, 1983.
- Idrus, Alwi. Perbandingan Kepekaan Model Mantel Haenszel dan Model Rasch dalam Mendeteksi Keberbedaan Fungsi Butir Ditinjau Dari Ukuran Responden. *Disertasi*. Jakarta: UNJ, 2011.
- Kadir. *Statistika Terapan. Konsep, Contoh, dan Analisis Data dengan Program SPSS/LISREL dalam Penelitian*. Jakarta: PT. Raja Grafindo Persada, 2015.
- Kartowagiran, B. Perbandingan Berbagai Metode untuk Mendeteksi Bias Butir. *Disertasi*. Yogyakarta: UGM, 2005.
- Keeves, J. K. (Ed), *The IEA Technical Handbook*. Hague: The International Association for The Evaluation of Educational Achievement, 1992.
- Kim, S., dan Lee, W. C. *IRT Scale Linking Methods for Mixed-Format Tests*. Iowa: ACT Inc., 2004.

Lampiran Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 53 Tahun 2015 tentang Penilaian Hasil Belajar oleh Pendidik pada Pendidikan Dasar dan Pendidikan Menengah.

Lampiran Peraturan Badan Standar Nasional Pendidikan Nomor 0031/P/BSNP/III/2015 tentang Prosedur Operasional Standar Penyelenggaraan Ujian Nasional Tahun Pelajaran 2014/2015.

Likun Hou, Jimmy de la Torre, dan Ratna Nandakumar. Differential Item Functioning Assessment in Cognitive Diagnostic Modeling: Application of the Wald Test to Investigate DIF in the DINA Model. *Journal of Educational Measurement Spring 2014, Vol. 51, No. 1, pp. 98–125* (diakses 5 Maret 2016).

Lord, F. M. *Applications of Item Response Theory to Practical Testing Problem*. New Jersey: Lawrence Erlbaum Associates Publishers, 1980.

Mazor, K. M. *Using Logistic Regression and Mantel Haenszel With Multiple Ability Estimates to Detect Differential Item Functioning*. *Journal of Education Measurement*. 32 (2).

Mehrens, W. A. & Lehmann, I. J. *Measurement and Evaluation in Education and Psychology*. New York: Rinehart & Wiston Inc., 1973.

Mikha Agus Widiyanto. *Statistika Terapan. Konsep & Aplikasi SPSS/LISREL dalam Penelitian Pendidikan, Psikologi & Ilmu Sosial Lainnya*. Jakarta: PT. Elex Media Komputindo, 2013.

Miller T., Chahine S., & Ruth A. Childs, Detecting Differential Item Functioning and Differential Step Functioning Due to Differences that Should Matter. *Journal Practical Assessment, Research & Evaluation, Volume 15, Number 10, July 2010, ISSN 1531-7714*, (diakses 25 Oktober 2015).

Mislevy R. J. & Bock R. D. *BILOG 3: Item Analysis and Test Scoring with Binary Logistic Models. Sccond Edition*, 1990.

Naga, Dali S. *Pengantar Teori Sekor pada Pengukuran Pendidikan*. Jakarta: Besbats, 1992.

_____. *Probalilitas dan Sekor pada Hipotesis Statistika*. Jakarta: UPT Penerbitan Universitas Tarumanagara, 2008.

- _____. *Teori Sekor pada Pengukuran Mental*. Jakarta: PT. Nagarani Citrayasa, 2013.
- Neisser, U. et. al., *Intelligence: Knowns and Unknowns*. American Psychological Association, Inc., 1996.
- Osterlind, S. J., *Test Item Bias*. Beverly Hill: Sage Publication Inc, 1983.
- Peraturan Pemerintah Republik Indonesia Nomor 13 Tahun 2015 tentang Perubahan Kedua Atas Peraturan Pemerintah Republik Indonesia Nomor 19 Tahun 2005 tentang Standar Nasional Pendidikan (SNP), pasal 1.*
- Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 66 Tahun 2014 tentang Standar Penilaian Pendidikan.*
- Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 59 Tahun 2014 tentang Kurikulum 2013 Sekolah Menengah Atas/Madrasah Aliyah, Lampiran III.*
- Peter Baldwin. On Mean-Sigma Estimators and Bias. *British Journal of Mathematical and Statistical Psychology* (2013), 66, 277–289 (diakses 4 Maret 2016).
- Pui-Wa Lei, Shu-Ying Chen, dan Lan Yu. Comparing Methods of Assessing Differential Item Functioning in a Computerized Adaptive Testing Environment. *Journal of Educational Measurement Vol. 43, No. 3, 2006* (diakses 9 Oktober 2015).
- Puspendik Balitbang Kemdikbud. *Panduan Pengembangan dan Pemberdayaan Bank Soal di Daerah*. Jakarta: Puspendik, 2013.
- Retnawati, H. Estimasi Efisiensi Relatif Tes Berdasarkan Teori Responsi Butir dan teori Tes Klasik. *Disertasi*. Yogyakarta: Universitas Negeri Yogyakarta, 2008.
- _____. *Teori Responsi Butir dan Penerapannya Untuk Peneliti, Praktisi Pengukuran dan Pengujian, Mahasiswa Pascasarjana*. Yogyakarta: Nuha Medika, 2014.
- _____. Keberfungsian Butir Diferensial pada Perangkat Tes Seleksi Masuk SLTP Mata Pelajaran Matematika. *Tesis* (Yogyakarta: UNY, 2003).

- Romel A. Morales, Evaluation of Mathematics Achievement Test: A Comparison between CTT and IRT. *The International Journal of Educational and Psychological Assessment* April 2009, Vol. 1, Issue 1, pp. 19-26 (diakses 5 Maret 2016).
- Salehi, Mohammad dan Alireza Tayebi. Differential Item Functioning: Implications for Test Validation. *Journal of Language Teaching and Research*, Vol. 3, No. 1, pp. 84-92, January 2012 ISSN 1798-4769 (diakses 5 Maret 2016).
- Seonghoon Kim & Won-Chan Lee. *IRT Scale Linking Methods for Mixed-Format Tests*. Iowa: ACT, Inc, 2004.
- Seock-Ho Kim & Allan S. Cohen. A Comparison of Lord's Chi-Square, Raju's Area Measures, and the Likelihood Ratio Test on Detection of Differential Item Functioning. *Journal Applied of Measurement in Education* 8, 4, 291-312 (diakses 5 Maret 2016).
- Seonghoon Kim & Won-Chan Lee. An Extension of Four IRT Linking Methods for Mixed-Format Tests. *Journal of Educational Measurement Spring* 2006, Vol. 43, No. 1, pp. 53-76 (diakses 5 Maret 2016).
- Setia Iriyanto dan Eny Winaryati, *Perbedaan Persepsi Antar Jenis Kelamin terhadap Peran Gender Dalam Keluarga Dan Masyarakat : Antara Harapan dan Kenyataan pada Guru-guru SD di Wilayah Kecamatan Tembalang Kota Semarang*. Prosiding Seminar Nasional Unimus. ISBN 978.979.704.883.9, 2010.
- Stark, S., & Chernyshenko, O. *Detection of Differentiation Item/Test Functioning (DIF/DTF) Using IRT*. Diakses 5 Agustus 2015, dari <http://www.work.psych.uiuc.edu/irt/>.
- Taehoon Kang & Nancy Petersen. *Linking Item Parameters to a Base Scale*. Iowa: ACT. Inc, 2009.
- Tim Penyusun. *Rancangan Penilaian*. Jakarta: Direktorat Pembinaan SMA, 2010.
- Walpole, R. E., Mers, R. H., Myers, S. R. *Probability and Statistic for Engineers and Scientist*. Upper Saddle River. NJ: Prentice Hall, 2002.

- Wardani Rahayu. Pengaruh Metode Linking terhadap Banyak Butir False Positive pada Pendeteksian DIF Berdasarkan Teori Responsi Butir. *Disertasi*. Jakarta: UNJ, 2007.
- Wiberg, Marie. Differential Item Functioning in Mastery Tests: A Comparison of Three Methods Using Real Data. *International Journal of Testing*, 9: 41–59, 2009, Taylor & Francis Group, LLC, ISSN: 1530-5058, (diakses 9 Oktober 2015).
- Won-Chan Lee & Jae-Chun Ban, *Comparison of Three IRT Linking Procedures in the Random Groups Equating Design*. Iowa: CASMA, 2007.
- Won-Chan Lee & Jae-Chun Ban, A Comparison of IRT Linking Procedures. *Journal Applied Measurement In Education*, 23: 23–48, 2010 (diakses 5 Maret 2016).
- World Health Organization. *Integrating gender into HIV/AIDS programmes in the health sector Tool to improve responsiveness to women's needs*. Geneva: WHO Press, 2009.
- Yanmei Li, Allan S. Cohen, dan Robert A. Ibarra, Characteristics of Mathematics Items Associated With Gender DIF. *International Journal of Testing*, Vol. 4, No. 2, 2004, Lawrence Erlbaum Associates, Inc. (diakses 11 September 2015).
- Yeni Tri Asmaningtyas, “Kemampuan Matematika Laki-laki dan Perempuan”, <http://download.portalgaruda.org/article.php?article=115727&val=5278> (diakses 10 Oktober 2015).
- Youngsuk Suh & Daniel M. Bolt. A Nested Logit Approach for Investigating Distractors as Causes of Differential Item Functioning. *Journal of Educational Measurement Summer 2011*, Vol. 48, No. 2, (diakses 9 Oktober 2015).

Lampiran 1. Uji Kecocokan Model dan Parameter Butir (Model L3P)

ITEM	INTERCEPT S.E.	SLOPE S.E.	THRESHOLD S.E.	DISPERSN S.E.	ASYMPTOTE S.E.	CHISQ (PROB)	DF
0001	-0.294 0.154*	1.122 0.166*	0.262 0.111*	0.891 0.132*	0.209 0.046*	3.1 (0.7954)	6.0
0002	0.029 0.109*	0.347 0.058*	-0.084 0.319*	2.882 0.483*	0.141 0.066*	47.2 (0.0000)	8.0
0003	1.334 0.119*	0.911 0.113*	-1.463 0.186*	1.097 0.136*	0.168 0.076*	19.0 (0.0043)	6.0
0004	1.311 0.112*	0.913 0.108*	-1.436 0.174*	1.095 0.130*	0.148 0.069*	9.0 (0.1722)	6.0
0005	0.523 0.093*	1.520 0.179*	-0.344 0.078*	0.658 0.078*	0.129 0.039*	8.2 (0.1462)	5.0
0006	0.711 0.113*	1.070 0.149*	-0.665 0.155*	0.935 0.130*	0.211 0.069*	3.6 (0.7315)	6.0
0007	0.100 0.106*	0.936 0.121*	-0.107 0.121*	1.068 0.138*	0.132 0.051*	3.5 (0.7522)	6.0
0008	0.151 0.117*	0.318 0.056*	-0.473 0.407*	3.140 0.556*	0.172 0.078*	34.1 (0.0000)	8.0
0009	-0.352 0.176*	2.114 0.307*	0.167 0.068*	0.473 0.069*	0.252 0.035*	3.1 (0.7952)	6.0
0010	0.565 0.081*	1.338 0.154*	-0.422 0.076*	0.747 0.086*	0.078 0.034*	6.8 (0.2326)	5.0
0011	-0.120 0.180*	1.584 0.286*	0.076 0.105*	0.631 0.114*	0.367 0.047*	2.6 (0.8559)	6.0
0012	-0.203 0.139*	0.230 0.050*	0.881 0.560*	4.340 0.940*	0.165 0.074*	95.5 (0.0000)	8.0
0013	0.424 0.124*	0.139 0.031*	-3.059 1.202*	7.206 1.633*	0.224 0.098*	79.3 (0.0000)	8.0
0014	-3.358 1.361*	1.169 0.572*	2.872 0.852*	0.855 0.419*	0.394 0.022*	10.6 (0.2241)	8.0
0015	1.738 0.140*	0.947 0.110*	-1.835 0.199*	1.056 0.123*	0.171 0.079*	6.9 (0.2250)	5.0
0016	1.795 0.149*	1.116 0.134*	-1.608 0.163*	0.896 0.107*	0.157 0.073*	5.1 (0.4062)	5.0
0017	0.024	1.876	-0.013	0.533	0.112	4.8	5.0

	0.108*	0.220*	0.058*	0.062*	0.031*	(0.4407)	
0018	0.984	1.091	-0.903	0.917	0.101	17.7	6.0
	0.090*	0.110*	0.114*	0.092*	0.047*	(0.0071)	
0019	-0.850	2.347	0.362	0.426	0.252	9.3	6.0
	0.256*	0.418*	0.060*	0.076*	0.030*	(0.1540)	
0020	-0.539	1.335	0.404	0.749	0.179	3.8	6.0
	0.166*	0.204*	0.084*	0.115*	0.038*	(0.7040)	
0021	-0.606	2.313	0.262	0.432	0.189	3.6	5.0
	0.188*	0.328*	0.056*	0.061*	0.029*	(0.6078)	
0022	1.122	1.171	-0.958	0.854	0.223	6.1	6.0
	0.124*	0.165*	0.164*	0.121*	0.077*	(0.4165)	
0023	0.479	1.170	-0.409	0.855	0.309	12.8	6.0
	0.133*	0.198*	0.153*	0.145*	0.068*	(0.0452)	
0024	-0.099	1.012	0.098	0.988	0.082	38.5	6.0
	0.091*	0.131*	0.084*	0.128*	0.036*	(0.0000)	
0025	-0.933	2.100	0.444	0.476	0.094	9.4	6.0
	0.193*	0.310*	0.048*	0.070*	0.023*	(0.1527)	
0026	0.995	1.244	-0.800	0.804	0.167	4.4	5.0
	0.109*	0.162*	0.128*	0.104*	0.063*	(0.4892)	
0027	-0.493	2.365	0.208	0.423	0.200	5.6	5.0
	0.183*	0.334*	0.057*	0.060*	0.030*	(0.3499)	
0028	0.237	1.122	-0.211	0.891	0.082	11.8	6.0
	0.081*	0.125*	0.082*	0.099*	0.035*	(0.0672)	
0029	-0.328	1.467	0.224	0.682	0.322	3.0	7.0
	0.187*	0.235*	0.103*	0.109*	0.044*	(0.8851)	
0030	-0.004	1.221	0.003	0.819	0.203	4.2	6.0
	0.126*	0.170*	0.103*	0.114*	0.047*	(0.6466)	
0031	0.067	0.782	-0.086	1.279	0.215	28.8	7.0
	0.140*	0.137*	0.188*	0.224*	0.071*	(0.0002)	
0032	1.139	1.217	-0.936	0.821	0.094	19.4	5.0
	0.097*	0.130*	0.105*	0.088*	0.045*	(0.0017)	
0033	-0.266	2.346	0.113	0.426	0.256	8.6	6.0
	0.180*	0.424*	0.064*	0.077*	0.037*	(0.1993)	
0034	-0.181	2.428	0.074	0.412	0.177	10.6	5.0
	0.145*	0.320*	0.055*	0.054*	0.032*	(0.0586)	
0035	-0.650	2.477	0.263	0.404	0.160	3.3	5.0
	0.189*	0.354*	0.051*	0.058*	0.027*	(0.6505)	

0036	-0.329	1.440	0.228	0.694	0.148	1.9	5.0
	0.135*	0.213*	0.074*	0.103*	0.036*	(0.8669)	
0037	-0.764	2.422	0.315	0.413	0.252	4.3	6.0
	0.243*	0.431*	0.060*	0.074*	0.032*	(0.6367)	
0038	1.067	0.769	-1.388	1.300	0.181	5.2	6.0
	0.110*	0.102*	0.222*	0.173*	0.081*	(0.5181)	
0039	-0.125	1.365	0.092	0.732	0.272	9.8	7.0
	0.155*	0.190*	0.105*	0.102*	0.046*	(0.2020)	
0040	0.184	0.797	-0.230	1.254	0.255	14.9	7.0
	0.145*	0.141*	0.209*	0.222*	0.076*	(0.0367)	

* STANDARD ERROR

LARGEST CHANGE = 0.009

579.5 242.0
(0.0000)

PARAMETER	MEAN	STN DEV
ASYMPTOTE	0.191	0.075
SLOPE	1.341	0.642
LOG (SLOPE)	0.141	0.638
THRESHOLD	-0.252	0.926

Lampiran 2. Nilai Fungsi Informasi Butir (Model L3P)

ITEM	MAXIMUM INFORMATION STANDARD ERROR *	POINT OF MAX INFORMATION STANDARD ERROR *	MAXIMUM EFFECTIVENESS POINT OF MAX EFFECTIVENESS *	AVERAGE INFORMATION INDEX OF RELIABILITY *
0001	0.6097 0.1321*	0.4065 0.1008*	0.2297 0.2824*	0.3264 0.2461*
0002	0.0663 0.0197*	0.2673 0.2994*	0.0263 0.0458*	0.0600 0.0566*
0003	0.4349 0.1085*	-1.3114 0.1729*	0.1066 -0.7235*	0.1923 0.1613*
0004	0.4528 0.1088*	-1.2978 0.1632*	0.1122 -0.7156*	0.2022 0.1682*
0005	1.3024 0.2636*	-0.2695 0.0729*	0.5048 -0.2136*	0.6091 0.3785*
0006	0.5516 0.1164*	-0.5122 0.1405*	0.2018 -0.3370*	0.3105 0.2369*
0007	0.4907 0.1002*	0.0168 0.1123*	0.1957 0.0100*	0.3194 0.2421*
0008	0.0527 0.0167*	-0.0306 0.3735*	0.0210 -0.0046*	0.0487 0.0464*
0009	1.9889 0.4489*	0.2540 0.0617*	0.7709 0.2261*	0.6776 0.4039*
0010	1.1108 0.2329*	-0.3655 0.0721*	0.4217 -0.2702*	0.5689 0.3626*
0011	0.8849 0.2162*	0.2242 0.0909*	0.3458 0.1849*	0.3802 0.2755*
0012	0.0280 0.0108*	1.4727 0.5697*	0.0102 0.1282*	0.0243 0.0237*
0013	0.0091 0.0041*	-1.8316 1.0422*	0.0034 -0.0567*	0.0085 0.0084*
0014	0.4555 0.3477*	3.0821 0.9137*	0.0056 2.2460*	0.0080 0.0079*
0015	0.4665 0.1193*	-1.6869 0.1874*	0.0819 -0.9351*	0.1510 0.1312*

0016	0.6661 0.1655*	-1.4906 0.1540*	0.1278 -0.9584*	0.2118 0.1748*
0017	2.0463 0.4141*	0.0414 0.0556*	0.8157 0.0354*	0.8246 0.4519*
0018	0.7062 0.1340*	-0.8165 0.1091*	0.2266 -0.5286*	0.3578 0.2635*
0019	2.4537 0.6971*	0.4407 0.0545*	0.8961 0.4009*	0.7142 0.4166*
0020	0.9138 0.2189*	0.5124 0.0783*	0.3299 0.3897*	0.4174 0.2945*
0021	2.6890 0.6351*	0.3274 0.0521*	1.0219 0.2961*	0.8380 0.4559*
0022	0.6456 0.1435*	-0.8134 0.1493*	0.2044 -0.5626*	0.3046 0.2335*
0023	0.5449 0.1254*	-0.2291 0.1336*	0.2133 -0.1626*	0.3014 0.2316*
0024	0.6310 0.1418*	0.1752 0.0815*	0.2493 0.1096*	0.3924 0.2818*
0025	2.6530 0.6974*	0.4863 0.0472*	0.9538 0.4279*	0.8706 0.4654*
0026	0.8105 0.1720*	-0.6892 0.1188*	0.2725 -0.4932*	0.3883 0.2797*
0027	2.7525 0.6396*	0.2748 0.0531*	1.0610 0.2496*	0.8534 0.4605*
0028	0.7749 0.1533*	-0.1402 0.0785*	0.3071 -0.0938*	0.4592 0.3147*
0029	0.8334 0.1855*	0.3714 0.0909*	0.3146 0.2976*	0.3656 0.2677*
0030	0.7306 0.1541*	0.1328 0.0934*	0.2896 0.0963*	0.3939 0.2826*
0031	0.2921 0.0712*	0.1258 0.1671*	0.1160 0.0655*	0.2046 0.1698*
0032	0.8921 0.1806*	-0.8634 0.1005*	0.2741 -0.5996*	0.4066 0.2890*
0033	2.4325	0.1927	0.9541	0.7683

		0.6814*		0.0575*		0.1751*		0.4345*
0034		3.0342		0.1336		1.2005		0.9539
		0.6689*		0.0515*		0.1218*		0.4882*
0035		3.2595		0.3166		1.2420		0.9667
		0.8004*		0.0476*		0.2894*		0.4915*
0036		1.1289		0.3155		0.4332		0.5259
		0.2713*		0.0691*		0.2468*		0.3447*
0037		2.6134		0.3914		0.9720		0.7551
		0.7382*		0.0544*		0.3579*		0.4302*
0038		0.3022		-1.1977		0.0851		0.1634
		0.0771*		0.2041*		-0.5658*		0.1404*
0039		0.7982		0.2339		0.3117		0.3865
		0.1588*		0.0943*		0.1809*		0.2788*
0040		0.2815		0.0026		0.1123		0.1963
		0.0664*		0.1827*		0.0014*		0.1641*

Lampiran 3. Syntax BILOG-MG (Model L3P)

```
F.BLM
>COMMENTS
>GLOBAL  DFNAME='F.PRN', NPARAM=3, SAVE;
>SAVE    PARM='F.PAR', SCORE='F.SCO', COV='F.COV';
>LENGTH  NITEMS=29;
>INPUT   NTOTAL=29, NALT=5, NIDCHAR=4;
>ITEMS   INAMES=(BUTIR01(1)BUTIR29);
>TEST1   TNAME='UN', INUMBER=(1(1)29);
(4A1,8X,29A1)
>CALIB   NQPT=31, CYCLES=25, NEWTON=10, CRIT=0.001, ACCEL=0.0,
CHI=15, PLOT=1;
>SCORE   NOPRINT, RSCTYPE=4, INFO=2, POP;
```

Lampiran 4. Data Hasil Analisis Bias Butir pada 30 Replikasi

REPLIKASI KE-1.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok		Indeks Subkelompok		Maksimum	Kesimpulan
1	R-F	0.06856	R1-R2	0.02738	0.02812	Bias
		Non-Uniform	F1-F2	0.02812		
2	R-F	0.02129	R1-R2	0.01914	0.07339	Tidak Bias
		Non-Uniform	F1-F2	0.07339		
3	R-F	0.02995	R1-R2	0.04965	0.04965	Tidak Bias
		Non-Uniform	F1-F2	0.02603		
4	R-F	0.03216	R1-R2	0.03189	0.03189	Bias
		Non-Uniform	F1-F2	0.02829		
5	R-F	0.01116	R1-R2	0.03944	0.03944	Tidak Bias
		Non-Uniform	F1-F2	0.01881		
6	R-F	0.01770	R1-R2	0.05748	0.05991	Tidak Bias
		Non-Uniform	F1-F2	0.05991		
7	R-F	0.03183	R1-R2	0.09074	0.09074	Tidak Bias
		Uniform	F1-F2	0.05965		
8	R-F	0.01733	R1-R2	0.05756	0.05756	Tidak Bias
		Non-Uniform	F1-F2	0.05046		
9	R-F	0.03740	R1-R2	0.07922	0.07922	Tidak Bias
		Uniform	F1-F2	0.04280		
10	R-F	0.01562	R1-R2	0.00922	0.02295	Tidak Bias
		Non-Uniform	F1-F2	0.02295		
11	R-F	0.04178	R1-R2	0.03023	0.03656	Bias
		Non-Uniform	F1-F2	0.03656		
12	R-F	0.02506	R1-R2	0.04776	0.06131	Tidak Bias
		Uniform	F1-F2	0.06131		
13	R-F	0.08594	R1-R2	0.06370	0.07459	Bias
		Non-Uniform	F1-F2	0.07459		
14	R-F	0.04056	R1-R2	0.03617	0.03617	Bias
		Non-Uniform	F1-F2	0.02487		
15	R-F	0.04202	R1-R2	0.05218	0.06330	Tidak Bias
		Uniform	F1-F2	0.06330		
16	R-F	0.01982	R1-R2	0.04166	0.04166	Tidak Bias
		Uniform	F1-F2	0.01910		
17	R-F	0.03760	R1-R2	0.10644	0.10644	Tidak Bias
		Non-Uniform	F1-F2	0.05335		
18	R-F	0.04622	R1-R2	0.04101	0.04896	Tidak Bias
		Non-Uniform	F1-F2	0.04896		
19	R-F	0.04880	R1-R2	0.01655	0.03788	Bias
		Non-Uniform	F1-F2	0.03788		
20	R-F	0.00841	R1-R2	0.03764	0.03764	Tidak Bias
		Non-Uniform	F1-F2	0.03686		

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
21	R-F	0.03809	R1-R2	0.06030	0.06522	Tidak Bias
		Non-Uniform	F1-F2	0.06522		
22	R-F	0.03804	R1-R2	0.03674	0.03674	Bias
		Non-Uniform	F1-F2	0.01686		
23	R-F	0.04657	R1-R2	0.03854	0.03854	Bias
		Uniform	F1-F2	0.02913		
24	R-F	0.05430	R1-R2	0.03836	0.03836	Bias
		Non-Uniform	F1-F2	0.01403		
25	R-F	0.03905	R1-R2	0.06086	0.06086	Tidak Bias
		Non-Uniform	F1-F2	0.03173		
26	R-F	0.02929	R1-R2	0.04532	0.04532	Tidak Bias
		Non-Uniform	F1-F2	0.03719		
27	R-F	0.04255	R1-R2	0.07882	0.07882	Tidak Bias
		Non-Uniform	F1-F2	0.02521		
28	R-F	0.06595	R1-R2	0.02451	0.03238	Bias
		Non-Uniform	F1-F2	0.03238		
29	R-F	0.03398	R1-R2	0.02817	0.05895	Tidak Bias
		Non-Uniform	F1-F2	0.05895		
JUMLAH BIAS BUTIR					10	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	10.33167520	7.81472	Bias
2	2.523510937	7.81472	Tidak Bias
3	2.612914135	7.81472	Tidak Bias
4	2.539497966	7.81472	Tidak Bias
5	0.382826271	7.81472	Tidak Bias
6	0.694020764	7.81472	Tidak Bias
7	4.144986054	7.81472	Tidak Bias
8	0.834626898	7.81472	Tidak Bias
9	-1.733706341	7.81472	Tidak Bias
10	0.856376805	7.81472	Tidak Bias
11	7.586446743	7.81472	Tidak Bias
12	4.39346758	7.81472	Tidak Bias
13	20.57384709	7.81472	Bias
14	4.452071774	7.81472	Tidak Bias
15	4.974765027	7.81472	Tidak Bias
16	1.410858799	7.81472	Tidak Bias
17	3.249445654	7.81472	Tidak Bias
18	4.906007919	7.81472	Tidak Bias
19	9.731952409	7.81472	Bias
20	0.202353021	7.81472	Tidak Bias
21	5.476895902	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
22	4.380457416	7.81472	Tidak Bias
23	10.61780741	7.81472	Bias
24	9.678094565	7.81472	Bias
25	3.594671878	7.81472	Tidak Bias
26	2.287554219	7.81472	Tidak Bias
27	6.803836481	7.81472	Tidak Bias
28	7.166181902	7.81472	Tidak Bias
29	2.973326306	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			5

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.37059	0.15023	-2.46679	Bias
2	-0.08765	0.26189	-0.33466	Tidak Bias
3	-0.13247	0.11914	-1.11189	Tidak Bias
4	0.21263	0.18384	1.15661	Tidak Bias
5	0.07638	0.16759	0.45575	Tidak Bias
6	0.04676	0.08206	0.56987	Tidak Bias
7	0.08629	0.12233	0.70541	Tidak Bias
8	0.07718	0.14688	0.52546	Tidak Bias
9	0.50194	5.61008	0.08947	Tidak Bias
10	0.15116	0.29264	0.51654	Tidak Bias
11	0.28658	0.17966	1.59512	Tidak Bias
12	0.00625	0.08903	0.07018	Tidak Bias
13	-0.19920	0.06531	-3.05002	Bias
14	-0.18057	0.10150	-1.77909	Tidak Bias
15	-0.09929	0.07232	-1.37292	Tidak Bias
16	0.20869	0.29659	0.70365	Tidak Bias
17	-0.08634	0.05679	-1.52020	Tidak Bias
18	0.19935	0.18704	1.06579	Tidak Bias
19	-0.10599	0.07115	-1.48973	Tidak Bias
20	-0.01141	0.14564	-0.07831	Tidak Bias
21	-0.04717	0.10747	-0.43889	Tidak Bias
22	-0.19065	0.15071	-1.26503	Tidak Bias
23	-0.13744	0.09607	-1.43070	Tidak Bias
24	-0.19342	0.08022	-2.41105	Bias
25	-0.12221	0.07739	-1.57917	Tidak Bias
26	0.04824	0.10967	0.43984	Tidak Bias
27	-0.09812	0.07491	-1.30972	Tidak Bias
28	0.44584	0.29301	1.52156	Tidak Bias
29	-0.28479	0.17095	-1.66586	Tidak Bias
JUMLAH BIAS BUTIR				3

REPLIKASI KE-2.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.03002	R1-R2	0.02529	0.07266	Tidak Bias
		Non-Uniform	F1-F2	0.07266		
2	R-F	0.02901	R1-R2	0.05532	0.05532	Tidak Bias
		Non-Uniform	F1-F2	0.01342		
3	R-F	0.03385	R1-R2	0.03860	0.03860	Tidak Bias
		Non-Uniform	F1-F2	0.03597		
4	R-F	0.02070	R1-R2	0.02581	0.05278	Tidak Bias
		Non-Uniform	F1-F2	0.05278		
5	R-F	0.03021	R1-R2	0.03159	0.05561	Tidak Bias
		Non-Uniform	F1-F2	0.05561		
6	R-F	0.03257	R1-R2	0.06678	0.06678	Tidak Bias
		Uniform	F1-F2	0.04397		
7	R-F	0.03958	R1-R2	0.03326	0.04530	Tidak Bias
		Non-Uniform	F1-F2	0.04530		
8	R-F	0.02831	R1-R2	0.06226	0.06226	Tidak Bias
		Non-Uniform	F1-F2	0.03297		
9	R-F	0.05630	R1-R2	0.06976	0.06976	Tidak Bias
		Non-Uniform	F1-F2	0.03376		
10	R-F	0.03546	R1-R2	0.03886	0.04844	Tidak Bias
		Non-Uniform	F1-F2	0.04844		
11	R-F	0.03706	R1-R2	0.01923	0.02578	Bias
		Non-Uniform	F1-F2	0.02578		
12	R-F	0.04298	R1-R2	0.03210	0.03210	Bias
		Non-Uniform	F1-F2	0.01679		
13	R-F	0.03820	R1-R2	0.02643	0.03239	Bias
		Uniform	F1-F2	0.03239		
14	R-F	0.05782	R1-R2	0.05538	0.06123	Tidak Bias
		Non-Uniform	F1-F2	0.06123		
15	R-F	0.03627	R1-R2	0.05370	0.05370	Tidak Bias
		Non-Uniform	F1-F2	0.04747		
16	R-F	0.04561	R1-R2	0.03451	0.03451	Bias
		Non-Uniform	F1-F2	0.00823		
17	R-F	0.01636	R1-R2	0.07733	0.07733	Tidak Bias
		Non-Uniform	F1-F2	0.05796		
18	R-F	0.04431	R1-R2	0.02389	0.05330	Tidak Bias
		Uniform	F1-F2	0.05330		
19	R-F	0.00807	R1-R2	0.04923	0.04923	Tidak Bias
		Non-Uniform	F1-F2	0.04072		
20	R-F	0.02752	R1-R2	0.06970	0.06970	Tidak Bias
		Non-Uniform	F1-F2	0.06081		
21	R-F	0.04436	R1-R2	0.07408	0.07408	Tidak Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
22	R-F	Non-Uniform	F1-F2	0.05131	0.03756	Tidak Bias
		0.03138	R1-R2	0.03756		
23	R-F	Non-Uniform	F1-F2	0.01410	0.03707	Tidak Bias
		0.03011	R1-R2	0.03292		
24	R-F	Non-Uniform	F1-F2	0.03707	0.04337	Tidak Bias
		0.02967	R1-R2	0.03708		
25	R-F	Uniform	F1-F2	0.04337	0.05940	Bias
		0.07691	R1-R2	0.05940		
26	R-F	Non-Uniform	F1-F2	0.00683	0.01643	Bias
		0.03722	R1-R2	0.01643		
27	R-F	Non-Uniform	F1-F2	0.01605	0.04922	Tidak Bias
		0.01716	R1-R2	0.04922		
28	R-F	Non-Uniform	F1-F2	0.04431	0.04203	Tidak Bias
		0.03833	R1-R2	0.00193		
29	R-F	Uniform	F1-F2	0.04203	0.04952	Tidak Bias
		0.04078	R1-R2	0.04952		
JUMLAH BIAS BUTIR					6	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	2.73475012	7.81472	Tidak Bias
2	6.175437913	7.81472	Tidak Bias
3	6.040984976	7.81472	Tidak Bias
4	1.138015039	7.81472	Tidak Bias
5	2.902267176	7.81472	Tidak Bias
6	3.690293789	7.81472	Tidak Bias
7	6.338474516	7.81472	Tidak Bias
8	2.65038227	7.81472	Tidak Bias
9	4.438808223	7.81472	Tidak Bias
10	3.383110782	7.81472	Tidak Bias
11	6.007244522	7.81472	Tidak Bias
12	9.361847446	7.81472	Bias
13	3.849129547	7.81472	Tidak Bias
14	11.08666335	7.81472	Bias
15	4.277678926	7.81472	Tidak Bias
16	5.361311491	7.81472	Tidak Bias
17	0.863564966	7.81472	Tidak Bias
18	10.84915139	7.81472	Bias
19	0.179021979	7.81472	Tidak Bias
20	3.106521673	7.81472	Tidak Bias
21	3.958192574	7.81472	Tidak Bias
22	2.730122057	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	3.667484361	7.81472	Tidak Bias
24	7.016602829	7.81472	Tidak Bias
25	21.97784015	7.81472	Bias
26	4.388414983	7.81472	Tidak Bias
27	0.706160829	7.81472	Tidak Bias
28	3.657511145	7.81472	Tidak Bias
29	7.706679752	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			4

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.15223	0.13454	-1.13151	Tidak Bias
2	-0.00897	0.23060	-0.03889	Tidak Bias
3	-0.10156	0.10610	-0.95720	Tidak Bias
4	0.14772	0.17683	0.83541	Tidak Bias
5	0.08366	0.17991	0.46501	Tidak Bias
6	0.04317	0.07779	0.55487	Tidak Bias
7	0.09794	0.10414	0.94048	Tidak Bias
8	0.11670	0.13862	0.84185	Tidak Bias
9	-0.36033	1.37132	-0.26276	Tidak Bias
10	-0.54550	0.34961	-1.56029	Tidak Bias
11	0.12242	0.18204	0.67248	Tidak Bias
12	0.11210	0.07716	1.45291	Tidak Bias
13	0.00138	0.06794	0.02035	Tidak Bias
14	-0.15161	0.09124	-1.66161	Tidak Bias
15	0.02504	0.07723	0.32421	Tidak Bias
16	-0.30808	0.24472	-1.25890	Tidak Bias
17	0.04237	0.05517	0.76794	Tidak Bias
18	0.14411	0.16427	0.87726	Tidak Bias
19	0.01862	0.06794	0.27413	Tidak Bias
20	0.02623	0.12526	0.20944	Tidak Bias
21	0.15521	0.09643	1.60962	Tidak Bias
22	0.04912	0.15467	0.31756	Tidak Bias
23	0.09246	0.08563	1.07970	Tidak Bias
24	0.00545	0.07778	0.07001	Tidak Bias
25	0.13878	0.06762	2.05217	Bias
26	0.17038	0.09831	1.73304	Tidak Bias
27	-0.00838	0.07638	-0.10972	Tidak Bias
28	0.10507	0.28401	0.36996	Tidak Bias
29	-0.06127	0.17381	-0.35251	Tidak Bias
JUMLAH BIAS BUTIR				1

REPLIKASI KE-3.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.05041	R1-R2	0.04411	0.04411	Bias
		Non-Uniform	F1-F2	0.02483		
2	R-F	0.03499	R1-R2	0.01216	0.04901	Tidak Bias
		Uniform	F1-F2	0.04901		
3	R-F	0.03395	R1-R2	0.03930	0.03930	Tidak Bias
		Non-Uniform	F1-F2	0.01508		
4	R-F	0.01488	R1-R2	0.04264	0.04264	Tidak Bias
		Uniform	F1-F2	0.02858		
5	R-F	0.02759	R1-R2	0.02613	0.03649	Tidak Bias
		Non-Uniform	F1-F2	0.03649		
6	R-F	0.02712	R1-R2	0.07596	0.07596	Tidak Bias
		Uniform	F1-F2	0.02631		
7	R-F	0.04101	R1-R2	0.04405	0.04405	Tidak Bias
		Uniform	F1-F2	0.01722		
8	R-F	0.04283	R1-R2	0.07958	0.07958	Tidak Bias
		Non-Uniform	F1-F2	0.03481		
9	R-F	0.03939	R1-R2	0.05155	0.05155	Tidak Bias
		Non-Uniform	F1-F2	0.03063		
10	R-F	0.02776	R1-R2	0.03541	0.03541	Tidak Bias
		Non-Uniform	F1-F2	0.01012		
11	R-F	0.02001	R1-R2	0.03154	0.03154	Tidak Bias
		Non-Uniform	F1-F2	0.02487		
12	R-F	0.02742	R1-R2	0.04001	0.04001	Tidak Bias
		Non-Uniform	F1-F2	0.01769		
13	R-F	0.04948	R1-R2	0.04482	0.04482	Bias
		Uniform	F1-F2	0.03766		
14	R-F	0.03567	R1-R2	0.08659	0.08659	Tidak Bias
		Non-Uniform	F1-F2	0.05787		
15	R-F	0.03274	R1-R2	0.04989	0.05455	Tidak Bias
		Non-Uniform	F1-F2	0.05455		
16	R-F	0.03511	R1-R2	0.01346	0.02129	Bias
		Non-Uniform	F1-F2	0.02129		
17	R-F	0.01583	R1-R2	0.06072	0.06072	Tidak Bias
		Non-Uniform	F1-F2	0.01395		
18	R-F	0.04440	R1-R2	0.03091	0.03091	Bias
		Uniform	F1-F2	0.00732		
19	R-F	0.02092	R1-R2	0.02881	0.02881	Tidak Bias
		Non-Uniform	F1-F2	0.02632		
20	R-F	0.02421	R1-R2	0.03449	0.03449	Tidak Bias
		Non-Uniform	F1-F2	0.02476		
21	R-F	0.05398	R1-R2	0.02628	0.02628	Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Non-Uniform	F1-F2	0.01412	Tidak Bias
		0.01623	R1-R2	0.06479	
23	R-F	Non-Uniform	F1-F2	0.03138	Bias
		0.04352	R1-R2	0.02599	
24	R-F	Uniform	F1-F2	0.02697	Tidak Bias
		0.02516	R1-R2	0.06464	
25	R-F	Non-Uniform	F1-F2	0.03561	Tidak Bias
		0.01152	R1-R2	0.09833	
26	R-F	Uniform	F1-F2	0.03264	Tidak Bias
		0.03371	R1-R2	0.03451	
27	R-F	Non-Uniform	F1-F2	0.03061	Tidak Bias
		0.01460	R1-R2	0.02624	
28	R-F	Uniform	F1-F2	0.01452	Bias
		0.04158	R1-R2	0.03802	
29	R-F	Non-Uniform	F1-F2	0.04148	Tidak Bias
		0.01363	R1-R2	0.03384	
JUMLAH BIAS BUTIR				7	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	6.869619001	7.81472	Tidak Bias
2	7.091944406	7.81472	Tidak Bias
3	2.611572166	7.81472	Tidak Bias
4	0.870469917	7.81472	Tidak Bias
5	2.241446622	7.81472	Tidak Bias
6	3.541346148	7.81472	Tidak Bias
7	6.643808808	7.81472	Tidak Bias
8	4.341845197	7.81472	Tidak Bias
9	1.430204421	7.81472	Tidak Bias
10	2.501727039	7.81472	Tidak Bias
11	1.475923655	7.81472	Tidak Bias
12	2.795106133	7.81472	Tidak Bias
13	13.07119136	7.81472	Bias
14	4.434689207	7.81472	Tidak Bias
15	2.895715327	7.81472	Tidak Bias
16	5.429297418	7.81472	Tidak Bias
17	1.081836009	7.81472	Tidak Bias
18	8.344884239	7.81472	Bias
19	1.283017676	7.81472	Tidak Bias
20	2.37422672	7.81472	Tidak Bias
21	5.528003049	7.81472	Tidak Bias
22	0.811937771	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	9.128120968	7.81472	Bias
24	2.277469167	7.81472	Tidak Bias
25	0.694515255	7.81472	Tidak Bias
26	3.033057172	7.81472	Tidak Bias
27	0.965376283	7.81472	Tidak Bias
28	2.626364958	7.81472	Tidak Bias
29	0.661485568	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			3

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.30497	0.13660	-2.23257	Bias
2	0.15648	0.24232	0.64575	Tidak Bias
3	-0.16786	0.11073	-1.51602	Tidak Bias
4	0.08268	0.17053	0.48481	Tidak Bias
5	0.04631	0.19389	0.23884	Tidak Bias
6	0.09499	0.08746	1.08619	Tidak Bias
7	0.14636	0.11799	1.24048	Tidak Bias
8	0.29931	0.15187	1.97076	Bias
9	-0.11374	2.03188	-0.05598	Tidak Bias
10	-0.43228	0.34163	-1.26535	Tidak Bias
11	0.09870	0.15884	0.62135	Tidak Bias
12	0.00580	0.08312	0.06980	Tidak Bias
13	-0.05526	0.06758	-0.81763	Tidak Bias
14	-0.11262	0.09290	-1.21229	Tidak Bias
15	-0.02532	0.08129	-0.31153	Tidak Bias
16	-0.20422	0.24334	-0.83925	Tidak Bias
17	0.01461	0.05815	0.25125	Tidak Bias
18	0.27430	0.17571	1.56107	Tidak Bias
19	0.02809	0.07021	0.40007	Tidak Bias
20	0.02846	0.12923	0.22020	Tidak Bias
21	0.10283	0.11205	0.91769	Tidak Bias
22	0.06390	0.14554	0.43905	Tidak Bias
23	-0.08086	0.09497	-0.85143	Tidak Bias
24	-0.07172	0.07838	-0.91501	Tidak Bias
25	-0.04198	0.07076	-0.59328	Tidak Bias
26	0.17008	0.10219	1.66437	Tidak Bias
27	-0.02807	0.08127	-0.34541	Tidak Bias
28	0.15998	0.28935	0.55289	Tidak Bias
29	-0.13395	0.19564	-0.68469	Tidak Bias
JUMLAH BIAS BUTIR				2

REPLIKASI KE-4.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok		Indeks Subkelompok		Maksimum	Kesimpulan
1	R-F	0.05627	R1-R2	0.04297	0.04297	Bias
		Non-Uniform	F1-F2	0.01039		
2	R-F	0.03838	R1-R2	0.03516	0.04231	Tidak Bias
		Non-Uniform	F1-F2	0.04231		
3	R-F	0.05552	R1-R2	0.03231	0.07090	Tidak Bias
		Non-Uniform	F1-F2	0.07090		
4	R-F	0.00638	R1-R2	0.03593	0.03593	Tidak Bias
		Non-Uniform	F1-F2	0.00741		
5	R-F	0.02993	R1-R2	0.04849	0.04849	Tidak Bias
		Non-Uniform	F1-F2	0.03372		
6	R-F	0.03552	R1-R2	0.03081	0.03081	Bias
		Non-Uniform	F1-F2	0.02082		
7	R-F	0.05942	R1-R2	0.04765	0.04765	Bias
		Non-Uniform	F1-F2	0.02402		
8	R-F	0.04313	R1-R2	0.04948	0.04948	Tidak Bias
		Non-Uniform	F1-F2	0.03485		
9	R-F	0.02642	R1-R2	0.02449	0.02449	Bias
		Uniform	F1-F2	0.00748		
10	R-F	0.01650	R1-R2	0.03648	0.03648	Tidak Bias
		Non-Uniform	F1-F2	0.00876		
11	R-F	0.03311	R1-R2	0.02535	0.04255	Tidak Bias
		Non-Uniform	F1-F2	0.04255		
12	R-F	0.03489	R1-R2	0.03412	0.03503	Tidak Bias
		Non-Uniform	F1-F2	0.03503		
13	R-F	0.04011	R1-R2	0.07944	0.07944	Tidak Bias
		Uniform	F1-F2	0.04875		
14	R-F	0.05704	R1-R2	0.03788	0.03801	Bias
		Non-Uniform	F1-F2	0.03801		
15	R-F	0.01918	R1-R2	0.02980	0.02980	Tidak Bias
		Non-Uniform	F1-F2	0.01238		
16	R-F	0.03888	R1-R2	0.06964	0.06964	Tidak Bias
		Non-Uniform	F1-F2	0.00766		
17	R-F	0.02605	R1-R2	0.03897	0.03897	Tidak Bias
		Non-Uniform	F1-F2	0.02108		
18	R-F	0.04092	R1-R2	0.04529	0.04529	Tidak Bias
		Uniform	F1-F2	0.02316		
19	R-F	0.01023	R1-R2	0.01942	0.06078	Tidak Bias
		Non-Uniform	F1-F2	0.06078		
20	R-F	0.03177	R1-R2	0.07932	0.07932	Tidak Bias
		Non-Uniform	F1-F2	0.05456		
21	R-F	0.04271	R1-R2	0.06251	0.06251	Tidak Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Uniform	F1-F2	0.04634	Tidak Bias
		0.01659	R1-R2	0.04170	
23	R-F	Uniform	F1-F2	0.00990	Bias
		0.02112	R1-R2	0.01422	
24	R-F	Non-Uniform	F1-F2	0.01191	Tidak Bias
		0.02621	R1-R2	0.07413	
25	R-F	Non-Uniform	F1-F2	0.04424	Tidak Bias
		0.01798	R1-R2	0.06792	
26	R-F	Non-Uniform	F1-F2	0.05694	Tidak Bias
		0.02723	R1-R2	0.03685	
27	R-F	Non-Uniform	F1-F2	0.04695	Tidak Bias
		0.03852	R1-R2	0.03123	
28	R-F	Non-Uniform	F1-F2	0.06303	Bias
		0.04118	R1-R2	0.01089	
29	R-F	Uniform	F1-F2	0.03737	Tidak Bias
		0.03790	R1-R2	0.04667	
Jumlah Bias Butir				7	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	10.86373436	7.81472	Bias
2	8.317385873	7.81472	Bias
3	6.084622699	7.81472	Tidak Bias
4	0.274251106	7.81472	Tidak Bias
5	2.756639355	7.81472	Tidak Bias
6	4.565972129	7.81472	Tidak Bias
7	10.19373897	7.81472	Bias
8	5.442576785	7.81472	Tidak Bias
9	1.377725189	7.81472	Tidak Bias
10	1.716687632	7.81472	Tidak Bias
11	28.69203935	7.81472	Bias
12	3.171289947	7.81472	Tidak Bias
13	5.551098644	7.81472	Tidak Bias
14	13.49476537	7.81472	Bias
15	1.031946056	7.81472	Tidak Bias
16	7.138206521	7.81472	Tidak Bias
17	2.062845573	7.81472	Tidak Bias
18	6.957917092	7.81472	Tidak Bias
19	0.366791925	7.81472	Tidak Bias
20	3.050698676	7.81472	Tidak Bias
21	4.861198563	7.81472	Tidak Bias
22	1.298546829	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	0.857279784	7.81472	Tidak Bias
24	1.259874106	7.81472	Tidak Bias
25	1.537381227	7.81472	Tidak Bias
26	1.636849385	7.81472	Tidak Bias
27	3.29885913	7.81472	Tidak Bias
28	4.459980124	7.81472	Tidak Bias
29	5.514794304	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			5

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.38117	0.12160	-3.13458	Bias
2	-0.04182	0.24081	-0.17365	Tidak Bias
3	-0.23855	0.12478	-1.91177	Tidak Bias
4	-0.04738	0.17116	-0.27683	Tidak Bias
5	0.07083	0.18356	0.38585	Tidak Bias
6	0.13869	0.08271	1.67687	Tidak Bias
7	0.21914	0.11105	1.97326	Bias
8	0.12646	0.17896	0.70665	Tidak Bias
9	-0.11670	2.00068	-0.05833	Tidak Bias
10	-0.33106	0.34612	-0.95648	Tidak Bias
11	0.15178	0.15562	0.97535	Tidak Bias
12	0.13187	0.07960	1.65664	Tidak Bias
13	-0.01918	0.07257	-0.26436	Tidak Bias
14	-0.30182	0.09829	-3.07060	Bias
15	0.00451	0.07730	0.05832	Tidak Bias
16	-0.23001	0.27767	-0.82838	Tidak Bias
17	0.02919	0.05825	0.50107	Tidak Bias
18	0.22492	0.18813	1.19556	Tidak Bias
19	0.02306	0.07298	0.31597	Tidak Bias
20	0.01508	0.13487	0.11182	Tidak Bias
21	0.12626	0.10313	1.22428	Tidak Bias
22	0.04531	0.16620	0.27265	Tidak Bias
23	0.05433	0.09267	0.58632	Tidak Bias
24	-0.04384	0.07451	-0.58837	Tidak Bias
25	0.06402	0.06911	0.92635	Tidak Bias
26	0.02388	0.10537	0.22666	Tidak Bias
27	0.08056	0.08338	0.96622	Tidak Bias
28	0.27433	0.28154	0.97440	Tidak Bias
29	-0.05272	0.17764	-0.29680	Tidak Bias
JUMLAH BIAS BUTIR				3

REPLIKASI KE-5.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.04456	R1-R2	0.05487	0.05487	Tidak Bias
		Non-Uniform	F1-F2	0.05042		
2	R-F	0.01059	R1-R2	0.04613	0.06530	Tidak Bias
		Uniform	F1-F2	0.06530		
3	R-F	0.02139	R1-R2	0.01649	0.03738	Tidak Bias
		Non-Uniform	F1-F2	0.03738		
4	R-F	0.01523	R1-R2	0.02285	0.02592	Tidak Bias
		Non-Uniform	F1-F2	0.02592		
5	R-F	0.01733	R1-R2	0.02522	0.06325	Tidak Bias
		Uniform	F1-F2	0.06325		
6	R-F	0.03748	R1-R2	0.01512	0.05288	Tidak Bias
		Non-Uniform	F1-F2	0.05288		
7	R-F	0.03852	R1-R2	0.03830	0.03830	Bias
		Uniform	F1-F2	0.03693		
8	R-F	0.04069	R1-R2	0.02782	0.04931	Tidak Bias
		Non-Uniform	F1-F2	0.04931		
9	R-F	0.04419	R1-R2	0.03998	0.03998	Bias
		Uniform	F1-F2	0.03598		
10	R-F	0.01389	R1-R2	0.02457	0.02457	Tidak Bias
		Non-Uniform	F1-F2	0.01448		
11	R-F	0.03987	R1-R2	0.03740	0.03740	Bias
		Non-Uniform	F1-F2	0.01437		
12	R-F	0.01702	R1-R2	0.03443	0.03541	Tidak Bias
		Non-Uniform	F1-F2	0.03541		
13	R-F	0.01511	R1-R2	0.02298	0.02298	Tidak Bias
		Non-Uniform	F1-F2	0.02169		
14	R-F	0.04542	R1-R2	0.04058	0.04451	Bias
		Non-Uniform	F1-F2	0.04451		
15	R-F	0.02234	R1-R2	0.03333	0.05574	Tidak Bias
		Non-Uniform	F1-F2	0.05574		
16	R-F	0.02855	R1-R2	0.04902	0.04902	Tidak Bias
		Non-Uniform	F1-F2	0.01643		
17	R-F	0.02287	R1-R2	0.08223	0.08223	Tidak Bias
		Non-Uniform	F1-F2	0.04954		
18	R-F	0.01617	R1-R2	0.00942	0.03474	Tidak Bias
		Non-Uniform	F1-F2	0.03474		
19	R-F	0.04905	R1-R2	0.03626	0.06641	Tidak Bias
		Non-Uniform	F1-F2	0.06641		
20	R-F	0.03826	R1-R2	0.06106	0.06106	Tidak Bias
		Non-Uniform	F1-F2	0.02672		
21	R-F	0.02326	R1-R2	0.02409	0.04488	Tidak Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Non-Uniform	F1-F2	0.04488	Tidak Bias
		0.03577	R1-R2	0.07755	
23	R-F	Non-Uniform	F1-F2	0.03135	Tidak Bias
		0.03837	R1-R2	0.04652	
24	R-F	Non-Uniform	F1-F2	0.06845	Tidak Bias
		0.02629	R1-R2	0.05590	
25	R-F	Non-Uniform	F1-F2	0.01950	Tidak Bias
		0.03039	R1-R2	0.02412	
26	R-F	Non-Uniform	F1-F2	0.03681	Tidak Bias
		0.01906	R1-R2	0.07837	
27	R-F	Non-Uniform	F1-F2	0.03876	Tidak Bias
		0.02787	R1-R2	0.03357	
28	R-F	Non-Uniform	F1-F2	0.04868	Bias
		0.07941	R1-R2	0.01666	
29	R-F	Non-Uniform	F1-F2	0.04817	Tidak Bias
		0.01576	R1-R2	0.02397	
JUMLAH BIAS BUTIR				5	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	5.197996225	7.81472	Tidak Bias
2	0.493393191	7.81472	Tidak Bias
3	1.033810542	7.81472	Tidak Bias
4	0.53318049	7.81472	Tidak Bias
5	1.612647807	7.81472	Tidak Bias
6	2.812322414	7.81472	Tidak Bias
7	6.213704869	7.81472	Tidak Bias
8	4.275493632	7.81472	Tidak Bias
9	4.441381165	7.81472	Tidak Bias
10	1.50194949	7.81472	Tidak Bias
11	5.226815223	7.81472	Tidak Bias
12	1.388360118	7.81472	Tidak Bias
13	0.926376544	7.81472	Tidak Bias
14	7.421201549	7.81472	Tidak Bias
15	1.762021934	7.81472	Tidak Bias
16	1.814375938	7.81472	Tidak Bias
17	1.004498031	7.81472	Tidak Bias
18	0.962301766	7.81472	Tidak Bias
19	10.54487335	7.81472	Bias
20	5.833235325	7.81472	Tidak Bias
21	1.335963248	7.81472	Tidak Bias
22	2.954299548	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	8.04178849	7.81472	Bias
24	1.729276417	7.81472	Tidak Bias
25	3.020411248	7.81472	Tidak Bias
26	0.987952102	7.81472	Tidak Bias
27	3.243235988	7.81472	Tidak Bias
28	8.719457608	7.81472	Bias
29	0.503262843	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			3

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.20348	0.11896	-1.71042	Tidak Bias
2	0.05041	0.26235	0.19214	Tidak Bias
3	-0.11175	0.13156	-0.84942	Tidak Bias
4	0.09690	0.16540	0.58586	Tidak Bias
5	0.06708	0.18509	0.36244	Tidak Bias
6	0.01027	0.07858	0.13066	Tidak Bias
7	0.10955	0.11503	0.95242	Tidak Bias
8	0.11635	0.16544	0.70327	Tidak Bias
9	0.17501	1.66886	0.10487	Tidak Bias
10	-0.23614	0.30238	-0.78094	Tidak Bias
11	0.18423	0.18871	0.97629	Tidak Bias
12	0.05745	0.07328	0.78387	Tidak Bias
13	-0.03320	0.07424	-0.44718	Tidak Bias
14	-0.25415	0.10730	-2.36866	Bias
15	-0.08088	0.07596	-1.06483	Tidak Bias
16	-0.17643	0.23368	-0.75501	Tidak Bias
17	-0.05816	0.05986	-0.97157	Tidak Bias
18	0.06047	0.19626	0.30811	Tidak Bias
19	-0.12027	0.07092	-1.69590	Tidak Bias
20	0.13073	0.12583	1.03900	Tidak Bias
21	0.12123	0.10984	1.10361	Tidak Bias
22	-0.04715	0.16794	-0.28076	Tidak Bias
23	-0.08897	0.08462	-1.05130	Tidak Bias
24	-0.05748	0.07937	-0.72418	Tidak Bias
25	-0.05204	0.06955	-0.74825	Tidak Bias
26	-0.02213	0.10875	-0.20345	Tidak Bias
27	-0.10152	0.07924	-1.28118	Tidak Bias
28	0.55667	0.28413	1.95918	Tidak Bias
29	-0.09260	0.17169	-0.53934	Tidak Bias
JUMLAH BIAS BUTIR				1

REPLIKASI KE-6.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.07385	R1-R2	0.05075	0.05075	Bias
		Non-Uniform	F1-F2	0.03751		
2	R-F	0.02670	R1-R2	0.02457	0.04086	Tidak Bias
		Uniform	F1-F2	0.04086		
3	R-F	0.02995	R1-R2	0.02959	0.03373	Tidak Bias
		Non-Uniform	F1-F2	0.03373		
4	R-F	0.03856	R1-R2	0.02659	0.02831	Bias
		Uniform	F1-F2	0.02831		
5	R-F	0.01632	R1-R2	0.01550	0.02322	Tidak Bias
		Non-Uniform	F1-F2	0.02322		
6	R-F	0.02990	R1-R2	0.07742	0.07742	Tidak Bias
		Non-Uniform	F1-F2	0.05501		
7	R-F	0.04127	R1-R2	0.02991	0.02991	Bias
		Uniform	F1-F2	0.02295		
8	R-F	0.02741	R1-R2	0.02976	0.04688	Tidak Bias
		Non-Uniform	F1-F2	0.04688		
9	R-F	0.01643	R1-R2	0.03808	0.03808	Tidak Bias
		Non-Uniform	F1-F2	0.01715		
10	R-F	0.00844	R1-R2	0.02106	0.02106	Tidak Bias
		Non-Uniform	F1-F2	0.01604		
11	R-F	0.05142	R1-R2	0.01458	0.01458	Bias
		Non-Uniform	F1-F2	0.00593		
12	R-F	0.02989	R1-R2	0.04002	0.05561	Tidak Bias
		Non-Uniform	F1-F2	0.05561		
13	R-F	0.05812	R1-R2	0.10045	0.10045	Tidak Bias
		Non-Uniform	F1-F2	0.07406		
14	R-F	0.03070	R1-R2	0.06054	0.06054	Tidak Bias
		Non-Uniform	F1-F2	0.05572		
15	R-F	0.05639	R1-R2	0.06008	0.06008	Tidak Bias
		Non-Uniform	F1-F2	0.02529		
16	R-F	0.01916	R1-R2	0.01754	0.04909	Tidak Bias
		Non-Uniform	F1-F2	0.04909		
17	R-F	0.03752	R1-R2	0.03488	0.03609	Bias
		Non-Uniform	F1-F2	0.03609		
18	R-F	0.01345	R1-R2	0.03226	0.03226	Tidak Bias
		Uniform	F1-F2	0.02181		
19	R-F	0.04758	R1-R2	0.07917	0.07917	Tidak Bias
		Non-Uniform	F1-F2	0.04717		
20	R-F	0.02347	R1-R2	0.02121	0.02427	Tidak Bias
		Non-Uniform	F1-F2	0.02427		
21	R-F	0.03323	R1-R2	0.02886	0.04018	Tidak Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Non-Uniform	F1-F2	0.04018	Bias
		0.04432	R1-R2	0.04009	
23	R-F	Non-Uniform	F1-F2	0.04216	Tidak Bias
		0.03871	R1-R2	0.04807	
24	R-F	Non-Uniform	F1-F2	0.07000	Bias
		0.03412	R1-R2	0.01604	
25	R-F	Non-Uniform	F1-F2	0.01775	Tidak Bias
		0.02732	R1-R2	0.06275	
26	R-F	Uniform	F1-F2	0.03336	Tidak Bias
		0.03893	R1-R2	0.07379	
27	R-F	Non-Uniform	F1-F2	0.06741	Bias
		0.05641	R1-R2	0.03662	
28	R-F	Non-Uniform	F1-F2	0.02875	Tidak Bias
		0.03907	R1-R2	0.03430	
29	R-F	Non-Uniform	F1-F2	0.06622	Tidak Bias
		0.01894	R1-R2	0.06763	
JUMLAH BIAS BUTIR				8	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	17.11770864	7.81472	Bias
2	3.274448606	7.81472	Tidak Bias
3	3.816619948	7.81472	Tidak Bias
4	8.203521558	7.81472	Bias
5	0.616775595	7.81472	Tidak Bias
6	3.193352384	7.81472	Tidak Bias
7	6.164467301	7.81472	Tidak Bias
8	2.22120729	7.81472	Tidak Bias
9	0.724602655	7.81472	Tidak Bias
10	0.644117456	7.81472	Tidak Bias
11	-40.90060724	7.81472	Tidak Bias
12	2.51626567	7.81472	Tidak Bias
13	15.27396846	7.81472	Bias
14	2.834291486	7.81472	Tidak Bias
15	9.193420735	7.81472	Bias
16	1.826626616	7.81472	Tidak Bias
17	3.111468274	7.81472	Tidak Bias
18	1.11510861	7.81472	Tidak Bias
19	8.050680889	7.81472	Bias
20	1.172660551	7.81472	Tidak Bias
21	2.006822657	7.81472	Tidak Bias
22	4.590471169	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	5.980807097	7.81472	Tidak Bias
24	3.405648919	7.81472	Tidak Bias
25	1.809505668	7.81472	Tidak Bias
26	4.889373334	7.81472	Tidak Bias
27	8.232250292	7.81472	Bias
28	2.838522286	7.81472	Tidak Bias
29	1.143379101	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			6

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.41453	0.13379	-3.09834	Bias
2	0.15506	0.27203	0.57001	Tidak Bias
3	-0.23355	0.12436	-1.87798	Tidak Bias
4	0.28140	0.16309	1.72548	Tidak Bias
5	-0.04998	0.16220	-0.30812	Tidak Bias
6	-0.10890	0.08444	-1.28968	Tidak Bias
7	0.17463	0.11247	1.55260	Tidak Bias
8	0.16959	0.13345	1.27082	Tidak Bias
9	0.27299	1.75759	0.15532	Tidak Bias
10	-0.19581	0.37694	-0.51947	Tidak Bias
11	0.30028	0.17907	1.67687	Tidak Bias
12	0.00670	0.08549	0.07836	Tidak Bias
13	-0.06445	0.07065	-0.91222	Tidak Bias
14	-0.11441	0.10024	-1.14127	Tidak Bias
15	-0.14244	0.07442	-1.91408	Tidak Bias
16	-0.00611	0.26103	-0.02342	Tidak Bias
17	-0.07629	0.05970	-1.27781	Tidak Bias
18	-0.00578	0.18393	-0.03141	Tidak Bias
19	-0.13682	0.07432	-1.84085	Tidak Bias
20	0.05861	0.14490	0.40451	Tidak Bias
21	-0.00158	0.10292	-0.01537	Tidak Bias
22	0.24932	0.15563	1.60195	Tidak Bias
23	-0.14430	0.09473	-1.52332	Tidak Bias
24	-0.15162	0.08549	-1.77351	Tidak Bias
25	-0.02480	0.07104	-0.34905	Tidak Bias
26	0.16397	0.09604	1.70728	Tidak Bias
27	-0.05129	0.08096	-0.63356	Tidak Bias
28	0.25856	0.26363	0.98078	Tidak Bias
29	-0.16846	0.18265	-0.92227	Tidak Bias
JUMLAH BIAS BUTIR				1

REPLIKASI KE-7.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.04043	R1-R2	0.08577	0.08577	Tidak Bias
		Non-Uniform	F1-F2	0.06265		
2	R-F	0.03209	R1-R2	0.01878	0.01878	Bias
		Uniform	F1-F2	0.01131		
3	R-F	0.02883	R1-R2	0.01921	0.03479	Tidak Bias
		Non-Uniform	F1-F2	0.03479		
4	R-F	0.01563	R1-R2	0.03148	0.03148	Tidak Bias
		Non-Uniform	F1-F2	0.03015		
5	R-F	0.04587	R1-R2	0.06392	0.06922	Tidak Bias
		Non-Uniform	F1-F2	0.06922		
6	R-F	0.03728	R1-R2	0.01658	0.11478	Tidak Bias
		Non-Uniform	F1-F2	0.11478		
7	R-F	0.03130	R1-R2	0.01638	0.02429	Bias
		Non-Uniform	F1-F2	0.02429		
8	R-F	0.02959	R1-R2	0.12059	0.12059	Tidak Bias
		Non-Uniform	F1-F2	0.03535		
9	R-F	0.01761	R1-R2	0.01385	0.06180	Tidak Bias
		Non-Uniform	F1-F2	0.06180		
10	R-F	0.02011	R1-R2	0.03059	0.03059	Tidak Bias
		Non-Uniform	F1-F2	0.02745		
11	R-F	0.00845	R1-R2	0.06333	0.06333	Tidak Bias
		Non-Uniform	F1-F2	0.05580		
12	R-F	0.04901	R1-R2	0.02605	0.02605	Bias
		Non-Uniform	F1-F2	0.02456		
13	R-F	0.08124	R1-R2	0.02071	0.08257	Tidak Bias
		Non-Uniform	F1-F2	0.08257		
14	R-F	0.04009	R1-R2	0.01888	0.06735	Tidak Bias
		Non-Uniform	F1-F2	0.06735		
15	R-F	0.02724	R1-R2	0.06081	0.06081	Tidak Bias
		Non-Uniform	F1-F2	0.04411		
16	R-F	0.03164	R1-R2	0.03779	0.03779	Tidak Bias
		Non-Uniform	F1-F2	0.01759		
17	R-F	0.04532	R1-R2	0.03759	0.09878	Tidak Bias
		Non-Uniform	F1-F2	0.09878		
18	R-F	0.05486	R1-R2	0.02353	0.07798	Tidak Bias
		Uniform	F1-F2	0.07798		
19	R-F	0.01138	R1-R2	0.02331	0.04802	Tidak Bias
		Non-Uniform	F1-F2	0.04802		
20	R-F	0.03713	R1-R2	0.00605	0.06595	Tidak Bias
		Non-Uniform	F1-F2	0.06595		
21	R-F	0.04770	R1-R2	0.02837	0.03390	Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Non-Uniform	F1-F2	0.03390	Bias
		0.04103	R1-R2	0.02289	
23	R-F	Non-Uniform	F1-F2	0.03280	Tidak Bias
		0.04708	R1-R2	0.09032	
24	R-F	Non-Uniform	F1-F2	0.06464	Tidak Bias
		0.03618	R1-R2	0.04270	
25	R-F	Non-Uniform	F1-F2	0.05405	Tidak Bias
		0.01688	R1-R2	0.02444	
26	R-F	Non-Uniform	F1-F2	0.09630	Tidak Bias
		0.02248	R1-R2	0.02465	
27	R-F	Non-Uniform	F1-F2	0.04911	Tidak Bias
		0.02900	R1-R2	0.01873	
28	R-F	Non-Uniform	F1-F2	0.08234	Tidak Bias
		0.04904	R1-R2	0.06562	
29	R-F	Non-Uniform	F1-F2	0.02547	Tidak Bias
		0.03728	R1-R2	0.06114	
JUMLAH BIAS BUTIR				5	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	5.490842642	7.81472	Tidak Bias
2	6.077438845	7.81472	Tidak Bias
3	2.98007124	7.81472	Tidak Bias
4	0.801578297	7.81472	Tidak Bias
5	5.241157237	7.81472	Tidak Bias
6	2.22483653	7.81472	Tidak Bias
7	3.712254773	7.81472	Tidak Bias
8	2.35205376	7.81472	Tidak Bias
9	0.872918198	7.81472	Tidak Bias
10	3.240203146	7.81472	Tidak Bias
11	-0.203841701	7.81472	Tidak Bias
12	5.013033419	7.81472	Tidak Bias
13	30.10847527	7.81472	Bias
14	4.237533767	7.81472	Tidak Bias
15	1.365063141	7.81472	Tidak Bias
16	3.804865945	7.81472	Tidak Bias
17	-21.5757384	7.81472	Tidak Bias
18	17.11082423	7.81472	Bias
19	0.596625937	7.81472	Tidak Bias
20	4.320342513	7.81472	Tidak Bias
21	4.481804074	7.81472	Tidak Bias
22	4.427135085	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	9.287809687	7.81472	Bias
24	3.14695001	7.81472	Tidak Bias
25	0.591113363	7.81472	Tidak Bias
26	1.207897426	7.81472	Tidak Bias
27	3.697901302	7.81472	Tidak Bias
28	3.648602635	7.81472	Tidak Bias
29	2.884701309	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			3

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.29879	0.14330	-2.08498	Bias
2	0.08440	0.25458	0.33151	Tidak Bias
3	-0.08646	0.11423	-0.75689	Tidak Bias
4	0.10306	0.15781	0.65305	Tidak Bias
5	-0.01421	0.18608	-0.07638	Tidak Bias
6	0.07506	0.08519	0.88108	Tidak Bias
7	0.08319	0.10202	0.81538	Tidak Bias
8	-0.01016	0.15970	-0.06361	Tidak Bias
9	-0.08556	1.52895	-0.05596	Tidak Bias
10	-0.38016	0.35918	-1.05841	Tidak Bias
11	0.00850	0.17650	0.04817	Tidak Bias
12	0.13884	0.08044	1.72589	Tidak Bias
13	-0.14444	0.07163	-2.01662	Bias
14	-0.08703	0.10492	-0.82946	Tidak Bias
15	-0.04837	0.08056	-0.60033	Tidak Bias
16	-0.24651	0.27455	-0.89786	Tidak Bias
17	0.06348	0.05513	1.15144	Tidak Bias
18	0.26730	0.18349	1.45673	Tidak Bias
19	-0.04011	0.07135	-0.56212	Tidak Bias
20	0.01572	0.12913	0.12177	Tidak Bias
21	0.09968	0.10788	0.92399	Tidak Bias
22	-0.02121	0.16081	-0.13189	Tidak Bias
23	-0.07859	0.09985	-0.78712	Tidak Bias
24	0.09633	0.07800	1.23504	Tidak Bias
25	0.04949	0.07116	0.69550	Tidak Bias
26	0.07623	0.10371	0.73500	Tidak Bias
27	-0.12744	0.07986	-1.59588	Tidak Bias
28	0.34109	0.29013	1.17564	Tidak Bias
29	0.16668	0.18978	0.87829	Tidak Bias
JUMLAH BIAS BUTIR				2

REPLIKASI KE-8.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.05089	R1-R2	0.05487	0.05487	Tidak Bias
		Non-Uniform	F1-F2	0.04157		
2	R-F	0.02892	R1-R2	0.04613	0.04613	Tidak Bias
		Uniform	F1-F2	0.03219		
3	R-F	0.02993	R1-R2	0.01649	0.01649	Bias
		Non-Uniform	F1-F2	0.01472		
4	R-F	0.02842	R1-R2	0.02285	0.04972	Tidak Bias
		Non-Uniform	F1-F2	0.04972		
5	R-F	0.00949	R1-R2	0.02522	0.04726	Tidak Bias
		Non-Uniform	F1-F2	0.04726		
6	R-F	0.04386	R1-R2	0.01512	0.03092	Bias
		Non-Uniform	F1-F2	0.03092		
7	R-F	0.02060	R1-R2	0.03830	0.03830	Tidak Bias
		Non-Uniform	F1-F2	0.02589		
8	R-F	0.00570	R1-R2	0.02782	0.02782	Tidak Bias
		Non-Uniform	F1-F2	0.02763		
9	R-F	0.05154	R1-R2	0.03998	0.03998	Bias
		Non-Uniform	F1-F2	0.03270		
10	R-F	0.01770	R1-R2	0.02457	0.02457	Tidak Bias
		Non-Uniform	F1-F2	0.00590		
11	R-F	0.05328	R1-R2	0.03740	0.03740	Bias
		Non-Uniform	F1-F2	0.02596		
12	R-F	0.03935	R1-R2	0.03443	0.05084	Tidak Bias
		Non-Uniform	F1-F2	0.05084		
13	R-F	0.06975	R1-R2	0.02298	0.07564	Tidak Bias
		Non-Uniform	F1-F2	0.07564		
14	R-F	0.08560	R1-R2	0.04058	0.04511	Bias
		Non-Uniform	F1-F2	0.04511		
15	R-F	0.06528	R1-R2	0.03333	0.04691	Bias
		Non-Uniform	F1-F2	0.04691		
16	R-F	0.03697	R1-R2	0.04902	0.04902	Tidak Bias
		Non-Uniform	F1-F2	0.03146		
17	R-F	0.05758	R1-R2	0.08223	0.08223	Tidak Bias
		Non-Uniform	F1-F2	0.04963		
18	R-F	0.00513	R1-R2	0.00942	0.02500	Tidak Bias
		Non-Uniform	F1-F2	0.02500		
19	R-F	0.05730	R1-R2	0.03626	0.05524	Bias
		Non-Uniform	F1-F2	0.05524		
20	R-F	0.00556	R1-R2	0.06106	0.06106	Tidak Bias
		Non-Uniform	F1-F2	0.00949		
21	R-F	0.04646	R1-R2	0.02409	0.03048	Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Non-Uniform	F1-F2	0.03048	Tidak Bias
		0.04383	R1-R2	0.07755	
23	R-F	Non-Uniform	F1-F2	0.01732	Bias
		0.05015	R1-R2	0.04652	
24	R-F	Non-Uniform	F1-F2	0.04633	Tidak Bias
		0.02564	R1-R2	0.05590	
25	R-F	Uniform	F1-F2	0.03170	Tidak Bias
		0.02469	R1-R2	0.02412	
26	R-F	Non-Uniform	F1-F2	0.02988	Tidak Bias
		0.03044	R1-R2	0.07837	
27	R-F	Non-Uniform	F1-F2	0.04018	Bias
		0.04809	R1-R2	0.03357	
28	R-F	Non-Uniform	F1-F2	0.02174	Bias
		0.07603	R1-R2	0.01666	
29	R-F	Non-Uniform	F1-F2	0.04058	Tidak Bias
		0.01428	R1-R2	0.02397	
JUMLAH BIAS BUTIR				11	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	9.074916045	7.81472	Bias
2	4.002778437	7.81472	Tidak Bias
3	2.067499491	7.81472	Tidak Bias
4	2.036190358	7.81472	Tidak Bias
5	0.459937643	7.81472	Tidak Bias
6	7.487774995	7.81472	Tidak Bias
7	0.959793043	7.81472	Tidak Bias
8	0.102944566	7.81472	Tidak Bias
9	7.080477378	7.81472	Tidak Bias
10	2.370644028	7.81472	Tidak Bias
11	14.27143925	7.81472	Bias
12	4.432145912	7.81472	Tidak Bias
13	19.28671493	7.81472	Bias
14	17.89691544	7.81472	Bias
15	12.07825619	7.81472	Bias
16	4.276493594	7.81472	Tidak Bias
17	7.729090916	7.81472	Tidak Bias
18	0.076850051	7.81472	Tidak Bias
19	13.56434303	7.81472	Bias
20	0.085647869	7.81472	Tidak Bias
21	4.26367133	7.81472	Tidak Bias
22	6.813901195	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	11.10339547	7.81472	Bias
24	4.879843078	7.81472	Tidak Bias
25	2.851611293	7.81472	Tidak Bias
26	2.383803351	7.81472	Tidak Bias
27	9.3344827	7.81472	Bias
28	9.383675129	7.81472	Bias
29	0.623811618	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			9

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.24969	0.12361	-2.01995	Bias
2	0.08765	0.24814	0.35322	Tidak Bias
3	-0.14357	0.12244	-1.17262	Tidak Bias
4	0.16371	0.15980	1.02447	Tidak Bias
5	0.09605	0.17614	0.54531	Tidak Bias
6	-0.11258	0.07479	-1.50517	Tidak Bias
7	0.05555	0.11109	0.50005	Tidak Bias
8	0.03585	0.17817	0.20122	Tidak Bias
9	0.70849	1.38154	0.51283	Tidak Bias
10	0.36658	0.35589	1.03004	Tidak Bias
11	0.29040	0.17109	1.69732	Tidak Bias
12	-0.01887	0.07857	-0.24010	Tidak Bias
13	-0.19760	0.07588	-2.60405	Bias
14	-0.41938	0.10099	-4.15256	Bias
15	-0.18518	0.07460	-2.48215	Bias
16	-0.27838	0.22597	-1.23194	Tidak Bias
17	-0.14049	0.05767	-2.43584	Bias
18	-0.02054	0.19901	-0.10321	Tidak Bias
19	-0.11676	0.07096	-1.64556	Tidak Bias
20	0.00539	0.13461	0.04005	Tidak Bias
21	0.08257	0.11843	0.69722	Tidak Bias
22	-0.19215	0.17901	-1.07338	Tidak Bias
23	-0.10683	0.08462	-1.26237	Tidak Bias
24	-0.06025	0.08042	-0.74916	Tidak Bias
25	-0.07305	0.06620	-1.10344	Tidak Bias
26	0.12942	0.10588	1.22233	Tidak Bias
27	-0.15350	0.07926	-1.93679	Tidak Bias
28	0.57416	0.28345	2.02560	Bias
29	-0.12703	0.16955	-0.74921	Tidak Bias
JUMLAH BIAS BUTIR				6

REPLIKASI KE-9.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.02310	R1-R2	0.02370	0.05517	Tidak Bias
		Non-Uniform	F1-F2	0.05517		
2	R-F	0.04249	R1-R2	0.03993	0.03993	Bias
		Non-Uniform	F1-F2	0.03462		
3	R-F	0.02394	R1-R2	0.01812	0.05552	Tidak Bias
		Non-Uniform	F1-F2	0.05552		
4	R-F	0.02490	R1-R2	0.03650	0.03650	Tidak Bias
		Non-Uniform	F1-F2	0.01857		
5	R-F	0.00753	R1-R2	0.04069	0.04069	Tidak Bias
		Non-Uniform	F1-F2	0.03331		
6	R-F	0.04730	R1-R2	0.03121	0.06365	Tidak Bias
		Non-Uniform	F1-F2	0.06365		
7	R-F	0.03585	R1-R2	0.00692	0.02568	Bias
		Uniform	F1-F2	0.02568		
8	R-F	0.04268	R1-R2	0.03087	0.05840	Tidak Bias
		Non-Uniform	F1-F2	0.05840		
9	R-F	0.03661	R1-R2	0.03307	0.03307	Bias
		Uniform	F1-F2	0.01957		
10	R-F	0.01398	R1-R2	0.01917	0.03733	Tidak Bias
		Non-Uniform	F1-F2	0.03733		
11	R-F	0.02292	R1-R2	0.04589	0.06342	Tidak Bias
		Non-Uniform	F1-F2	0.06342		
12	R-F	0.04316	R1-R2	0.04374	0.04455	Tidak Bias
		Non-Uniform	F1-F2	0.04455		
13	R-F	0.04546	R1-R2	0.06017	0.06017	Tidak Bias
		Non-Uniform	F1-F2	0.05856		
14	R-F	0.06498	R1-R2	0.03414	0.03414	Bias
		Non-Uniform	F1-F2	0.02258		
15	R-F	0.03323	R1-R2	0.03177	0.06781	Tidak Bias
		Non-Uniform	F1-F2	0.06781		
16	R-F	0.01163	R1-R2	0.01844	0.01844	Tidak Bias
		Non-Uniform	F1-F2	0.00581		
17	R-F	0.01704	R1-R2	0.00870	0.09297	Tidak Bias
		Non-Uniform	F1-F2	0.09297		
18	R-F	0.00979	R1-R2	0.03967	0.03967	Tidak Bias
		Non-Uniform	F1-F2	0.03267		
19	R-F	0.01977	R1-R2	0.01146	0.01146	Bias
		Non-Uniform	F1-F2	0.00848		
20	R-F	0.01169	R1-R2	0.02634	0.05389	Tidak Bias
		Uniform	F1-F2	0.05389		
21	R-F	0.04260	R1-R2	0.06262	0.06822	Tidak Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Non-Uniform	F1-F2	0.06822	Bias
		0.06354	R1-R2	0.02997	
		Non-Uniform	F1-F2	0.05721	
23	R-F	0.04522	R1-R2	0.04798	Tidak Bias
		Uniform	F1-F2	0.01799	
24	R-F	0.02792	R1-R2	0.00557	Tidak Bias
		Non-Uniform	F1-F2	0.04177	
25	R-F	0.02195	R1-R2	0.02129	Tidak Bias
		Non-Uniform	F1-F2	0.07402	
26	R-F	0.02022	R1-R2	0.06855	Tidak Bias
		Non-Uniform	F1-F2	0.11748	
27	R-F	0.02971	R1-R2	0.01767	Tidak Bias
		Non-Uniform	F1-F2	0.05042	
28	R-F	0.04757	R1-R2	0.02550	Bias
		Uniform	F1-F2	0.00480	
29	R-F	0.01162	R1-R2	0.03598	Tidak Bias
		Uniform	F1-F2	0.05216	
JUMLAH BIAS BUTIR				7	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	1.895303656	7.81472	Tidak Bias
2	10.33904852	7.81472	Bias
3	1.426253539	7.81472	Tidak Bias
4	2.113323256	7.81472	Tidak Bias
5	0.106619542	7.81472	Tidak Bias
6	8.595824562	7.81472	Bias
7	5.364038799	7.81472	Tidak Bias
8	4.858277458	7.81472	Tidak Bias
9	-35.23331116	7.81472	Tidak Bias
10	1.460546514	7.81472	Tidak Bias
11	7.309556906	7.81472	Tidak Bias
12	7.034854603	7.81472	Tidak Bias
13	3.984602519	7.81472	Tidak Bias
14	11.47517625	7.81472	Bias
15	1.941881742	7.81472	Tidak Bias
16	0.766526232	7.81472	Tidak Bias
17	0.696628824	7.81472	Tidak Bias
18	0.420232706	7.81472	Tidak Bias
19	1.92177646	7.81472	Tidak Bias
20	0.6286027	7.81472	Tidak Bias
21	3.130771336	7.81472	Tidak Bias
22	8.672345076	7.81472	Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	15.87920097	7.81472	Bias
24	1.415477864	7.81472	Tidak Bias
25	1.264631225	7.81472	Tidak Bias
26	0.753513902	7.81472	Tidak Bias
27	3.07552275	7.81472	Tidak Bias
28	6.481871895	7.81472	Tidak Bias
29	0.656605396	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			5

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.10487	0.16867	-0.62173	Tidak Bias
2	-0.12670	0.28222	-0.44894	Tidak Bias
3	-0.15360	0.13718	-1.11969	Tidak Bias
4	0.11517	0.14768	0.77987	Tidak Bias
5	-0.02578	0.17532	-0.14703	Tidak Bias
6	0.21282	0.07910	2.69048	Bias
7	0.10174	0.11512	0.88374	Tidak Bias
8	0.32382	0.16488	1.96396	Bias
9	0.05848	1.13264	0.05163	Tidak Bias
10	-0.02552	0.32044	-0.07964	Tidak Bias
11	0.00425	0.19431	0.02187	Tidak Bias
12	0.16971	0.07870	2.15636	Bias
13	0.02853	0.07142	0.39946	Tidak Bias
14	-0.35268	0.10676	-3.30358	Bias
15	-0.07943	0.07700	-1.03157	Tidak Bias
16	-0.01688	0.25163	-0.06708	Tidak Bias
17	-0.03805	0.05821	-0.65366	Tidak Bias
18	0.03144	0.20394	0.15417	Tidak Bias
19	-0.06543	0.07533	-0.86862	Tidak Bias
20	0.01640	0.13796	0.11887	Tidak Bias
21	0.15561	0.10724	1.45115	Tidak Bias
22	-0.32566	0.15878	-2.05109	Bias
23	0.02015	0.09074	0.22203	Tidak Bias
24	-0.08856	0.08037	-1.10194	Tidak Bias
25	0.00694	0.07138	0.09719	Tidak Bias
26	-0.08448	0.11955	-0.70661	Tidak Bias
27	-0.04596	0.07898	-0.58190	Tidak Bias
28	0.25922	0.26983	0.96068	Tidak Bias
29	0.02932	0.18047	0.16246	Tidak Bias
JUMLAH BIAS BUTIR				5

REPLIKASI KE-10.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.05990	R1-R2	0.05216	0.05216	Bias
		Non-Uniform	F1-F2	0.04204		
2	R-F	0.03217	R1-R2	0.00983	0.03166	Bias
		Uniform	F1-F2	0.03166		
3	R-F	0.01739	R1-R2	0.04269	0.04269	Tidak Bias
		Non-Uniform	F1-F2	0.01487		
4	R-F	0.02548	R1-R2	0.02387	0.04915	Tidak Bias
		Non-Uniform	F1-F2	0.04915		
5	R-F	0.01517	R1-R2	0.05839	0.05839	Tidak Bias
		Non-Uniform	F1-F2	0.04746		
6	R-F	0.04023	R1-R2	0.05882	0.05882	Tidak Bias
		Non-Uniform	F1-F2	0.03113		
7	R-F	0.03207	R1-R2	0.05601	0.05601	Tidak Bias
		Non-Uniform	F1-F2	0.02588		
8	R-F	0.01538	R1-R2	0.06355	0.06355	Tidak Bias
		Non-Uniform	F1-F2	0.02769		
9	R-F	0.05404	R1-R2	0.03510	0.03510	Bias
		Non-Uniform	F1-F2	0.03247		
10	R-F	0.01511	R1-R2	0.02650	0.02650	Tidak Bias
		Non-Uniform	F1-F2	0.00613		
11	R-F	0.06086	R1-R2	0.03483	0.03483	Bias
		Non-Uniform	F1-F2	0.02622		
12	R-F	0.03855	R1-R2	0.03142	0.05078	Tidak Bias
		Non-Uniform	F1-F2	0.05078		
13	R-F	0.08343	R1-R2	0.04422	0.07784	Bias
		Non-Uniform	F1-F2	0.07784		
14	R-F	0.08215	R1-R2	0.05979	0.05979	Bias
		Non-Uniform	F1-F2	0.04515		
15	R-F	0.06481	R1-R2	0.10491	0.10491	Tidak Bias
		Non-Uniform	F1-F2	0.04581		
16	R-F	0.04635	R1-R2	0.02123	0.03073	Bias
		Non-Uniform	F1-F2	0.03073		
17	R-F	0.05207	R1-R2	0.08365	0.08365	Tidak Bias
		Non-Uniform	F1-F2	0.05200		
18	R-F	0.01192	R1-R2	0.04279	0.04279	Tidak Bias
		Non-Uniform	F1-F2	0.02466		
19	R-F	0.04446	R1-R2	0.06440	0.06440	Tidak Bias
		Non-Uniform	F1-F2	0.05692		
20	R-F	0.01011	R1-R2	0.05456	0.05456	Tidak Bias
		Non-Uniform	F1-F2	0.00948		
21	R-F	0.02907	R1-R2	0.06210	0.06210	Tidak Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Non-Uniform	F1-F2	0.03103	Tidak Bias
		0.02442	R1-R2	0.05104	
23	R-F	Uniform	F1-F2	0.01752	Bias
		0.05044	R1-R2	0.03896	
24	R-F	Non-Uniform	F1-F2	0.04519	Tidak Bias
		0.03221	R1-R2	0.04360	
25	R-F	Uniform	F1-F2	0.03170	Tidak Bias
		0.01972	R1-R2	0.05226	
26	R-F	Non-Uniform	F1-F2	0.02916	Tidak Bias
		0.02615	R1-R2	0.04412	
27	R-F	Non-Uniform	F1-F2	0.04106	Tidak Bias
		0.04280	R1-R2	0.05707	
28	R-F	Uniform	F1-F2	0.02246	Bias
		0.05873	R1-R2	0.01091	
29	R-F	Non-Uniform	F1-F2	0.04030	Tidak Bias
		0.01721	R1-R2	0.02558	
JUMLAH BIAS BUTIR				9	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	12.43669914	7.81472	Bias
2	4.410014044	7.81472	Tidak Bias
3	1.170636761	7.81472	Tidak Bias
4	1.965565317	7.81472	Tidak Bias
5	0.714134597	7.81472	Tidak Bias
6	3.297162933	7.81472	Tidak Bias
7	2.244522228	7.81472	Tidak Bias
8	0.736178908	7.81472	Tidak Bias
9	7.48886837	7.81472	Tidak Bias
10	1.47586911	7.81472	Tidak Bias
11	35.63822302	7.81472	Bias
12	6.835413177	7.81472	Tidak Bias
13	30.81700625	7.81472	Bias
14	17.06711878	7.81472	Bias
15	12.04638363	7.81472	Bias
16	6.791992531	7.81472	Tidak Bias
17	7.454934244	7.81472	Tidak Bias
18	0.959211877	7.81472	Tidak Bias
19	7.249858873	7.81472	Tidak Bias
20	0.664302144	7.81472	Tidak Bias
21	1.526166272	7.81472	Tidak Bias
22	1.976366176	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	11.53877848	7.81472	Bias
24	6.368927669	7.81472	Tidak Bias
25	1.469009045	7.81472	Tidak Bias
26	1.672750413	7.81472	Tidak Bias
27	7.37669797	7.81472	Tidak Bias
28	6.256227787	7.81472	Tidak Bias
29	1.03715989	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			6

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.33875	0.13003	-2.60527	Bias
2	0.14251	0.27650	0.51541	Tidak Bias
3	-0.07486	0.12069	-0.62024	Tidak Bias
4	0.20349	0.17042	1.19405	Tidak Bias
5	0.08216	0.19061	0.43105	Tidak Bias
6	0.00515	0.07525	0.06842	Tidak Bias
7	0.07763	0.10936	0.70985	Tidak Bias
8	0.07305	0.17043	0.42864	Tidak Bias
9	0.71907	1.64248	0.43780	Tidak Bias
10	0.36451	0.37276	0.97787	Tidak Bias
11	0.39540	0.17974	2.19984	Bias
12	-0.11640	0.08004	-1.45439	Tidak Bias
13	-0.22028	0.07374	-2.98740	Bias
14	-0.38931	0.09750	-3.99299	Bias
15	-0.23839	0.07936	-3.00404	Bias
16	-0.33690	0.26409	-1.27571	Tidak Bias
17	-0.14490	0.05630	-2.57360	Bias
18	-0.12267	0.18100	-0.67773	Tidak Bias
19	-0.05359	0.07323	-0.73185	Tidak Bias
20	-0.04882	0.12660	-0.38563	Tidak Bias
21	0.10958	0.12758	0.85893	Tidak Bias
22	-0.05146	0.19131	-0.26900	Tidak Bias
23	-0.19983	0.09530	-2.09683	Bias
24	-0.10330	0.08523	-1.21190	Tidak Bias
25	-0.02791	0.06650	-0.41965	Tidak Bias
26	0.09018	0.10937	0.82453	Tidak Bias
27	-0.12644	0.07981	-1.58431	Tidak Bias
28	0.46653	0.29638	1.57409	Tidak Bias
29	-0.13545	0.17148	-0.78991	Tidak Bias
JUMLAH BIAS BUTIR				7

REPLIKASI KE-11.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.04437	R1-R2	0.09029	0.09029	Tidak Bias
		Non-Uniform	F1-F2	0.02110		
2	R-F	0.04677	R1-R2	0.02502	0.03440	Bias
		Non-Uniform	F1-F2	0.03440		
3	R-F	0.03085	R1-R2	0.02377	0.02975	Bias
		Non-Uniform	F1-F2	0.02975		
4	R-F	0.04228	R1-R2	0.00630	0.02289	Bias
		Non-Uniform	F1-F2	0.02289		
5	R-F	0.02832	R1-R2	0.03332	0.04364	Tidak Bias
		Non-Uniform	F1-F2	0.04364		
6	R-F	0.03497	R1-R2	0.07860	0.07860	Tidak Bias
		Uniform	F1-F2	0.03601		
7	R-F	0.04791	R1-R2	0.09042	0.09042	Tidak Bias
		Uniform	F1-F2	0.01820		
8	R-F	0.04373	R1-R2	0.04326	0.04326	Bias
		Non-Uniform	F1-F2	0.02582		
9	R-F	0.03831	R1-R2	0.01154	0.01154	Bias
		Non-Uniform	F1-F2	0.00564		
10	R-F	0.02666	R1-R2	0.02125	0.02716	Tidak Bias
		Non-Uniform	F1-F2	0.02716		
11	R-F	0.01121	R1-R2	0.01014	0.02024	Tidak Bias
		Uniform	F1-F2	0.02024		
12	R-F	0.03933	R1-R2	0.05095	0.05095	Tidak Bias
		Non-Uniform	F1-F2	0.04495		
13	R-F	0.04208	R1-R2	0.02454	0.06658	Tidak Bias
		Non-Uniform	F1-F2	0.06658		
14	R-F	0.03581	R1-R2	0.04886	0.04886	Tidak Bias
		Non-Uniform	F1-F2	0.02750		
15	R-F	0.03350	R1-R2	0.00592	0.01650	Bias
		Non-Uniform	F1-F2	0.01650		
16	R-F	0.02642	R1-R2	0.02200	0.03004	Tidak Bias
		Non-Uniform	F1-F2	0.03004		
17	R-F	0.03593	R1-R2	0.02255	0.04101	Tidak Bias
		Non-Uniform	F1-F2	0.04101		
18	R-F	0.02929	R1-R2	0.04231	0.04231	Tidak Bias
		Uniform	F1-F2	0.02498		
19	R-F	0.01150	R1-R2	0.04755	0.06376	Tidak Bias
		Uniform	F1-F2	0.06376		
20	R-F	0.03560	R1-R2	0.02961	0.02961	Bias
		Non-Uniform	F1-F2	0.02137		
21	R-F	0.02709	R1-R2	0.05712	0.06511	Tidak Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
22	R-F	Uniform	F1-F2	0.06511	0.02863	Tidak Bias
		0.01003	R1-R2	0.01594		
		Non-Uniform	F1-F2	0.02863		
23	R-F	0.01880	R1-R2	0.03543	0.03543	Tidak Bias
		Non-Uniform	F1-F2	0.01964		
		0.03238	R1-R2	0.03012		
24	R-F	Uniform	F1-F2	0.04013	0.04013	Tidak Bias
		0.02317	R1-R2	0.01293		
25	R-F	Non-Uniform	F1-F2	0.04902	0.04902	Tidak Bias
		0.03309	R1-R2	0.00944		
26	R-F	Non-Uniform	F1-F2	0.01986	0.01986	Bias
		0.02119	R1-R2	0.05474		
27	R-F	Non-Uniform	F1-F2	0.04412	0.05474	Tidak Bias
		0.05218	R1-R2	0.00931		
28	R-F	Non-Uniform	F1-F2	0.05453	0.05453	Tidak Bias
		0.02508	R1-R2	0.05505		
29	R-F	Non-Uniform	F1-F2	0.00915	0.05505	Tidak Bias
		JUMLAH BIAS BUTIR				

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	5.439684851	7.81472	Tidak Bias
2	13.97768606	7.81472	Bias
3	2.654233647	7.81472	Tidak Bias
4	7.064869348	7.81472	Tidak Bias
5	1.889491135	7.81472	Tidak Bias
6	4.922882819	7.81472	Tidak Bias
7	9.176779386	7.81472	Bias
8	6.287048324	7.81472	Tidak Bias
9	1.117762853	7.81472	Tidak Bias
10	6.724607791	7.81472	Tidak Bias
11	2.204339373	7.81472	Tidak Bias
12	10.3555672	7.81472	Bias
13	3.760780944	7.81472	Tidak Bias
14	2.70848746	7.81472	Tidak Bias
15	2.640108448	7.81472	Tidak Bias
16	2.86857985	7.81472	Tidak Bias
17	6.653994674	7.81472	Tidak Bias
18	5.957042284	7.81472	Tidak Bias
19	0.616582744	7.81472	Tidak Bias
20	4.886671114	7.81472	Tidak Bias
21	3.559482111	7.81472	Tidak Bias
22	0.213954045	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	1.979947411	7.81472	Tidak Bias
24	6.057825739	7.81472	Tidak Bias
25	1.059944695	7.81472	Tidak Bias
26	2.278844911	7.81472	Tidak Bias
27	1.645981998	7.81472	Tidak Bias
28	3.711057021	7.81472	Tidak Bias
29	2.801168122	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			3

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.27232	0.13301	-2.04741	Bias
2	-0.16997	0.27452	-0.61916	Tidak Bias
3	-0.11560	0.11889	-0.97232	Tidak Bias
4	0.21418	0.17428	1.22896	Tidak Bias
5	-0.07082	0.16838	-0.42059	Tidak Bias
6	0.10638	0.08108	1.31201	Tidak Bias
7	0.15310	0.11953	1.28080	Tidak Bias
8	0.25283	0.14118	1.79082	Tidak Bias
9	-0.36645	2.18012	-0.16809	Tidak Bias
10	-0.56637	0.39857	-1.42101	Tidak Bias
11	0.04380	0.18358	0.23858	Tidak Bias
12	0.06109	0.07734	0.78987	Tidak Bias
13	0.01157	0.07499	0.15427	Tidak Bias
14	0.03800	0.09468	0.40140	Tidak Bias
15	-0.01431	0.07928	-0.18049	Tidak Bias
16	-0.16698	0.25157	-0.66377	Tidak Bias
17	0.00272	0.05566	0.04887	Tidak Bias
18	0.06995	0.17539	0.39882	Tidak Bias
19	0.00572	0.07015	0.08148	Tidak Bias
20	0.06941	0.11375	0.61025	Tidak Bias
21	0.06282	0.11216	0.56004	Tidak Bias
22	-0.05738	0.15785	-0.36350	Tidak Bias
23	0.01608	0.09084	0.17700	Tidak Bias
24	0.04908	0.08715	0.56312	Tidak Bias
25	0.06319	0.06939	0.91059	Tidak Bias
26	0.15289	0.11271	1.35654	Tidak Bias
27	0.08466	0.08451	1.00175	Tidak Bias
28	0.32984	0.28715	1.14868	Tidak Bias
29	0.01290	0.17624	0.07320	Tidak Bias
JUMLAH BIAS BUTIR				1

REPLIKASI KE-12.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.04893	R1-R2	0.02832	0.06730	Tidak Bias
		Non-Uniform	F1-F2	0.06730		
2	R-F	0.03803	R1-R2	0.01714	0.02635	Bias
		Non-Uniform	F1-F2	0.02635		
3	R-F	0.06263	R1-R2	0.02750	0.04773	Bias
		Non-Uniform	F1-F2	0.04773		
4	R-F	0.01656	R1-R2	0.03189	0.06427	Tidak Bias
		Non-Uniform	F1-F2	0.06427		
5	R-F	0.02232	R1-R2	0.06062	0.06062	Tidak Bias
		Non-Uniform	F1-F2	0.03332		
6	R-F	0.01199	R1-R2	0.04047	0.04047	Tidak Bias
		Uniform	F1-F2	0.02179		
7	R-F	0.04694	R1-R2	0.01446	0.05931	Tidak Bias
		Non-Uniform	F1-F2	0.05931		
8	R-F	0.03070	R1-R2	0.03531	0.03660	Tidak Bias
		Non-Uniform	F1-F2	0.03660		
9	R-F	0.05323	R1-R2	0.07199	0.07199	Tidak Bias
		Uniform	F1-F2	0.01557		
10	R-F	0.01538	R1-R2	0.01408	0.01408	Bias
		Uniform	F1-F2	0.01031		
11	R-F	0.02012	R1-R2	0.01233	0.05577	Tidak Bias
		Non-Uniform	F1-F2	0.05577		
12	R-F	0.02395	R1-R2	0.04051	0.04051	Tidak Bias
		Uniform	F1-F2	0.01709		
13	R-F	0.03122	R1-R2	0.02831	0.03143	Tidak Bias
		Non-Uniform	F1-F2	0.03143		
14	R-F	0.03478	R1-R2	0.06592	0.07531	Tidak Bias
		Non-Uniform	F1-F2	0.07531		
15	R-F	0.02801	R1-R2	0.03368	0.09964	Tidak Bias
		Non-Uniform	F1-F2	0.09964		
16	R-F	0.04824	R1-R2	0.05318	0.05318	Tidak Bias
		Non-Uniform	F1-F2	0.02351		
17	R-F	0.02305	R1-R2	0.03331	0.06650	Tidak Bias
		Non-Uniform	F1-F2	0.06650		
18	R-F	0.03360	R1-R2	0.02684	0.05263	Tidak Bias
		Non-Uniform	F1-F2	0.05263		
19	R-F	0.02792	R1-R2	0.04978	0.04978	Tidak Bias
		Non-Uniform	F1-F2	0.03587		
20	R-F	0.03542	R1-R2	0.04176	0.04176	Tidak Bias
		Non-Uniform	F1-F2	0.00772		
21	R-F	0.04560	R1-R2	0.04117	0.06436	Tidak Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
22	R-F	Uniform	F1-F2	0.06436	0.06019	Tidak Bias
		0.02107	R1-R2	0.03201		
		Non-Uniform	F1-F2	0.06019		
23	R-F	0.05079	R1-R2	0.07168	0.07168	Tidak Bias
		Non-Uniform	F1-F2	0.05614		
		0.02129	R1-R2	0.03022		
24	R-F	Uniform	F1-F2	0.04327	0.04327	Tidak Bias
		0.01575	R1-R2	0.07982		
25	R-F	Non-Uniform	F1-F2	0.07590	0.07982	Tidak Bias
		0.03235	R1-R2	0.03744		
26	R-F	Non-Uniform	F1-F2	0.03234	0.03744	Tidak Bias
		0.01978	R1-R2	0.05105		
27	R-F	Non-Uniform	F1-F2	0.05221	0.05221	Tidak Bias
		0.02254	R1-R2	0.03125		
28	R-F	Non-Uniform	F1-F2	0.01812	0.03125	Tidak Bias
		0.02921	R1-R2	0.05690		
29	R-F	Uniform	F1-F2	0.05875	0.05875	Tidak Bias
		JUMLAH BIAS BUTIR				

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	7.201283513	7.81472	Tidak Bias
2	7.34662048	7.81472	Tidak Bias
3	7.937030621	7.81472	Bias
4	0.92652569	7.81472	Tidak Bias
5	1.770691349	7.81472	Tidak Bias
6	0.455561301	7.81472	Tidak Bias
7	7.42841284	7.81472	Tidak Bias
8	2.079952928	7.81472	Tidak Bias
9	5.061878303	7.81472	Tidak Bias
10	1.975415394	7.81472	Tidak Bias
11	1.321638633	7.81472	Tidak Bias
12	2.821854284	7.81472	Tidak Bias
13	3.368376463	7.81472	Tidak Bias
14	4.72119406	7.81472	Tidak Bias
15	2.110090363	7.81472	Tidak Bias
16	7.573539723	7.81472	Tidak Bias
17	4.9970253	7.81472	Tidak Bias
18	4.546899245	7.81472	Tidak Bias
19	3.305114055	7.81472	Tidak Bias
20	4.548740959	7.81472	Tidak Bias
21	6.237677819	7.81472	Tidak Bias
22	1.030992332	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	8.077668176	7.81472	Bias
24	2.782868763	7.81472	Tidak Bias
25	0.893668195	7.81472	Tidak Bias
26	2.953680067	7.81472	Tidak Bias
27	1.178603023	7.81472	Tidak Bias
28	1.428588904	7.81472	Tidak Bias
29	3.513358564	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			2

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.27194	0.14074	-1.93225	Tidak Bias
2	0.14637	0.24066	0.60821	Tidak Bias
3	-0.27722	0.12830	-2.16079	Bias
4	0.08299	0.15209	0.54568	Tidak Bias
5	0.11930	0.17439	0.68411	Tidak Bias
6	0.00785	0.08399	0.09348	Tidak Bias
7	0.16961	0.11120	1.52527	Tidak Bias
8	-0.01244	0.16857	-0.07377	Tidak Bias
9	0.08901	1.11557	0.07979	Tidak Bias
10	0.06407	0.36748	0.17434	Tidak Bias
11	-0.04314	0.17080	-0.25255	Tidak Bias
12	0.03462	0.07868	0.44001	Tidak Bias
13	-0.02926	0.06867	-0.42612	Tidak Bias
14	-0.18141	0.10078	-1.80008	Tidak Bias
15	-0.07673	0.07766	-0.98803	Tidak Bias
16	-0.37342	0.27402	-1.36273	Tidak Bias
17	0.01742	0.05507	0.31638	Tidak Bias
18	0.19290	0.17051	1.13130	Tidak Bias
19	-0.02549	0.07472	-0.34119	Tidak Bias
20	0.13242	0.14122	0.93769	Tidak Bias
21	0.10391	0.10288	1.01002	Tidak Bias
22	-0.00228	0.17228	-0.01322	Tidak Bias
23	-0.06083	0.08996	-0.67617	Tidak Bias
24	0.05872	0.08123	0.72295	Tidak Bias
25	0.00404	0.07688	0.05253	Tidak Bias
26	0.07717	0.11274	0.68456	Tidak Bias
27	-0.00459	0.08022	-0.05726	Tidak Bias
28	0.16091	0.28648	0.56168	Tidak Bias
29	-0.10259	0.19005	-0.53981	Tidak Bias
JUMLAH BIAS BUTIR				1

REPLIKASI KE-13.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.04360	R1-R2	0.01820	0.06146	Tidak Bias
		Non-Uniform	F1-F2	0.06146		
2	R-F	0.02757	R1-R2	0.06965	0.06965	Tidak Bias
		Non-Uniform	F1-F2	0.01154		
3	R-F	0.02906	R1-R2	0.01735	0.04353	Tidak Bias
		Non-Uniform	F1-F2	0.04353		
4	R-F	0.02517	R1-R2	0.02970	0.04105	Tidak Bias
		Non-Uniform	F1-F2	0.04105		
5	R-F	0.02561	R1-R2	0.02229	0.04494	Tidak Bias
		Non-Uniform	F1-F2	0.04494		
6	R-F	0.04805	R1-R2	0.06479	0.06479	Tidak Bias
		Non-Uniform	F1-F2	0.04248		
7	R-F	0.02222	R1-R2	0.01643	0.01643	Bias
		Uniform	F1-F2	0.01441		
8	R-F	0.04376	R1-R2	0.05664	0.05664	Tidak Bias
		Non-Uniform	F1-F2	0.03864		
9	R-F	0.05502	R1-R2	0.04434	0.04434	Bias
		Uniform	F1-F2	0.03534		
10	R-F	0.00680	R1-R2	0.01679	0.03038	Tidak Bias
		Non-Uniform	F1-F2	0.03038		
11	R-F	0.01923	R1-R2	0.06250	0.06762	Tidak Bias
		Non-Uniform	F1-F2	0.06762		
12	R-F	0.02109	R1-R2	0.05830	0.05830	Tidak Bias
		Non-Uniform	F1-F2	0.01635		
13	R-F	0.05702	R1-R2	0.09179	0.09179	Tidak Bias
		Non-Uniform	F1-F2	0.01873		
14	R-F	0.03240	R1-R2	0.07376	0.07376	Tidak Bias
		Non-Uniform	F1-F2	0.04526		
15	R-F	0.02890	R1-R2	0.09939	0.09939	Tidak Bias
		Non-Uniform	F1-F2	0.02784		
16	R-F	0.06234	R1-R2	0.04081	0.04081	Bias
		Non-Uniform	F1-F2	0.01855		
17	R-F	0.04368	R1-R2	0.07344	0.07344	Tidak Bias
		Non-Uniform	F1-F2	0.01480		
18	R-F	0.01977	R1-R2	0.02601	0.06053	Tidak Bias
		Non-Uniform	F1-F2	0.06053		
19	R-F	0.00814	R1-R2	0.07204	0.07204	Tidak Bias
		Non-Uniform	F1-F2	0.03723		
20	R-F	0.02764	R1-R2	0.02896	0.02896	Tidak Bias
		Non-Uniform	F1-F2	0.02582		
21	R-F	0.05348	R1-R2	0.09916	0.09916	Tidak Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
22	R-F	Non-Uniform	F1-F2	0.02163	0.04245	Tidak Bias
		0.02707	R1-R2	0.04245		
23	R-F	Non-Uniform	F1-F2	0.03886	0.06808	Tidak Bias
		0.02802	R1-R2	0.06808		
24	R-F	Non-Uniform	F1-F2	0.02355	0.04334	Bias
		0.04589	R1-R2	0.04334		
25	R-F	Non-Uniform	F1-F2	0.01732	0.05634	Tidak Bias
		0.04505	R1-R2	0.05634		
26	R-F	Non-Uniform	F1-F2	0.03064	0.05985	Tidak Bias
		0.03874	R1-R2	0.05985		
27	R-F	Uniform	F1-F2	0.02453	0.07684	Tidak Bias
		0.01717	R1-R2	0.07684		
28	R-F	Non-Uniform	F1-F2	0.07318	0.05578	Tidak Bias
		0.04355	R1-R2	0.05578		
29	R-F	Uniform	F1-F2	0.03417	0.03241	Tidak Bias
		0.02108	R1-R2	0.03241		
JUMLAH BIAS BUTIR					4	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	5.790985505	7.81472	Tidak Bias
2	5.564346012	7.81472	Tidak Bias
3	2.374706384	7.81472	Tidak Bias
4	2.191655499	7.81472	Tidak Bias
5	1.512871888	7.81472	Tidak Bias
6	8.227443508	7.81472	Bias
7	2.271939011	7.81472	Tidak Bias
8	4.611947516	7.81472	Tidak Bias
9	4.298680329	7.81472	Tidak Bias
10	0.531040215	7.81472	Tidak Bias
11	1.224654795	7.81472	Tidak Bias
12	2.2394572	7.81472	Tidak Bias
13	6.55550046	7.81472	Tidak Bias
14	3.24306633	7.81472	Tidak Bias
15	2.647218594	7.81472	Tidak Bias
16	9.297975063	7.81472	Bias
17	4.02093151	7.81472	Tidak Bias
18	1.697916529	7.81472	Tidak Bias
19	0.221983114	7.81472	Tidak Bias
20	2.162874192	7.81472	Tidak Bias
21	6.945952144	7.81472	Tidak Bias
22	1.515641719	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	2.012042233	7.81472	Tidak Bias
24	8.209029199	7.81472	Bias
25	7.769401169	7.81472	Tidak Bias
26	4.507359167	7.81472	Tidak Bias
27	0.975759314	7.81472	Tidak Bias
28	4.536528417	7.81472	Tidak Bias
29	1.258931672	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			3

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.35768	0.15978	-2.23858	Bias
2	0.00706	0.24605	0.02870	Tidak Bias
3	-0.19385	0.12962	-1.49556	Tidak Bias
4	0.12241	0.19636	0.62341	Tidak Bias
5	-0.00420	0.17816	-0.02357	Tidak Bias
6	0.10686	0.07714	1.38515	Tidak Bias
7	0.06050	0.10487	0.57693	Tidak Bias
8	0.24169	0.13401	1.80352	Tidak Bias
9	-0.07243	1.92080	-0.03771	Tidak Bias
10	-0.01906	0.37376	-0.05100	Tidak Bias
11	0.07655	0.18227	0.41997	Tidak Bias
12	0.04190	0.07942	0.52762	Tidak Bias
13	0.02490	0.07220	0.34490	Tidak Bias
14	-0.09133	0.10559	-0.86492	Tidak Bias
15	-0.06915	0.07600	-0.90989	Tidak Bias
16	-0.38447	0.24959	-1.54039	Tidak Bias
17	0.09236	0.05883	1.56980	Tidak Bias
18	-0.03630	0.19446	-0.18668	Tidak Bias
19	0.02918	0.07051	0.41391	Tidak Bias
20	-0.00725	0.13729	-0.05282	Tidak Bias
21	0.31166	0.12238	2.54668	Bias
22	-0.10144	0.14986	-0.67686	Tidak Bias
23	0.01999	0.10450	0.19131	Tidak Bias
24	-0.09105	0.08779	-1.03708	Tidak Bias
25	0.10037	0.07753	1.29460	Tidak Bias
26	0.09386	0.11389	0.82415	Tidak Bias
27	-0.03670	0.07726	-0.47506	Tidak Bias
28	0.26852	0.26238	1.02339	Tidak Bias
29	-0.13292	0.20574	-0.64602	Tidak Bias
JUMLAH BIAS BUTIR				2

REPLIKASI KE-14.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.05808	R1-R2	0.05216	0.05216	Bias
		Non-Uniform	F1-F2	0.04047		
2	R-F	0.01862	R1-R2	0.00983	0.03389	Tidak Bias
		Uniform	F1-F2	0.03389		
3	R-F	0.01011	R1-R2	0.04269	0.04269	Tidak Bias
		Non-Uniform	F1-F2	0.02381		
4	R-F	0.03022	R1-R2	0.02387	0.09444	Tidak Bias
		Non-Uniform	F1-F2	0.09444		
5	R-F	0.00594	R1-R2	0.05839	0.08526	Tidak Bias
		Non-Uniform	F1-F2	0.08526		
6	R-F	0.05702	R1-R2	0.05882	0.05882	Tidak Bias
		Non-Uniform	F1-F2	0.02991		
7	R-F	0.03684	R1-R2	0.05601	0.05601	Tidak Bias
		Non-Uniform	F1-F2	0.03752		
8	R-F	0.02389	R1-R2	0.06355	0.06355	Tidak Bias
		Non-Uniform	F1-F2	0.03461		
9	R-F	0.05275	R1-R2	0.03510	0.03771	Bias
		Uniform	F1-F2	0.03771		
10	R-F	0.01162	R1-R2	0.02650	0.03272	Tidak Bias
		Non-Uniform	F1-F2	0.03272		
11	R-F	0.05262	R1-R2	0.03483	0.04405	Bias
		Non-Uniform	F1-F2	0.04405		
12	R-F	0.02733	R1-R2	0.03142	0.05348	Tidak Bias
		Non-Uniform	F1-F2	0.05348		
13	R-F	0.06197	R1-R2	0.04422	0.05584	Bias
		Non-Uniform	F1-F2	0.05584		
14	R-F	0.07499	R1-R2	0.05979	0.10642	Tidak Bias
		Non-Uniform	F1-F2	0.10642		
15	R-F	0.04712	R1-R2	0.10491	0.10491	Tidak Bias
		Non-Uniform	F1-F2	0.05533		
16	R-F	0.02761	R1-R2	0.02123	0.02123	Bias
		Non-Uniform	F1-F2	0.01922		
17	R-F	0.05385	R1-R2	0.08365	0.08365	Tidak Bias
		Non-Uniform	F1-F2	0.03668		
18	R-F	0.02268	R1-R2	0.04279	0.04279	Tidak Bias
		Non-Uniform	F1-F2	0.02884		
19	R-F	0.03908	R1-R2	0.06440	0.06440	Tidak Bias
		Uniform	F1-F2	0.03262		
20	R-F	0.01287	R1-R2	0.05456	0.05456	Tidak Bias
		Uniform	F1-F2	0.03113		
21	R-F	0.02680	R1-R2	0.06210	0.07481	Tidak Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Non-Uniform	F1-F2	0.07481	Tidak Bias
		0.03376	R1-R2	0.05104	
23	R-F	Uniform	F1-F2	0.03401	Tidak Bias
		0.02633	R1-R2	0.03896	
24	R-F	Non-Uniform	F1-F2	0.02791	Bias
		0.04758	R1-R2	0.04360	
25	R-F	Non-Uniform	F1-F2	0.03578	Tidak Bias
		0.01172	R1-R2	0.05226	
26	R-F	Non-Uniform	F1-F2	0.05633	Tidak Bias
		0.02208	R1-R2	0.04412	
27	R-F	Non-Uniform	F1-F2	0.04230	Tidak Bias
		0.00887	R1-R2	0.05707	
28	R-F	Non-Uniform	F1-F2	0.05484	Bias
		0.03208	R1-R2	0.01091	
29	R-F	Uniform	F1-F2	0.02440	Tidak Bias
		0.02474	R1-R2	0.02558	
JUMLAH BIAS BUTIR				7	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	11.17587435	7.81472	Bias
2	1.416903626	7.81472	Tidak Bias
3	0.397035263	7.81472	Tidak Bias
4	2.361344099	7.81472	Tidak Bias
5	0.075643705	7.81472	Tidak Bias
6	5.722752501	7.81472	Tidak Bias
7	3.3484033	7.81472	Tidak Bias
8	1.988764054	7.81472	Tidak Bias
9	5.848766368	7.81472	Tidak Bias
10	1.383287727	7.81472	Tidak Bias
11	14.67332237	7.81472	Bias
12	4.003109002	7.81472	Tidak Bias
13	10.55830772	7.81472	Bias
14	14.08245148	7.81472	Bias
15	5.458617414	7.81472	Tidak Bias
16	2.582582188	7.81472	Tidak Bias
17	8.968198929	7.81472	Bias
18	1.639902168	7.81472	Tidak Bias
19	4.165359959	7.81472	Tidak Bias
20	0.938765752	7.81472	Tidak Bias
21	2.534804252	7.81472	Tidak Bias
22	3.979533566	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	3.580261837	7.81472	Tidak Bias
24	7.271046818	7.81472	Tidak Bias
25	0.445592492	7.81472	Tidak Bias
26	1.327554108	7.81472	Tidak Bias
27	0.280814718	7.81472	Tidak Bias
28	2.914633797	7.81472	Tidak Bias
29	1.870649433	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			5

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.33040	0.13089	-2.52428	Bias
2	0.14041	0.28851	0.48668	Tidak Bias
3	-0.06341	0.12089	-0.52454	Tidak Bias
4	0.16383	0.18371	0.89174	Tidak Bias
5	-0.00551	0.19384	-0.02844	Tidak Bias
6	0.05053	0.07838	0.64465	Tidak Bias
7	0.11638	0.10786	1.07895	Tidak Bias
8	0.16100	0.15744	1.02262	Tidak Bias
9	0.49391	1.66709	0.29627	Tidak Bias
10	0.16483	0.36335	0.45364	Tidak Bias
11	0.36593	0.17244	2.12212	Bias
12	-0.05843	0.07477	-0.78137	Tidak Bias
13	-0.14476	0.07141	-2.02708	Bias
14	-0.34373	0.10297	-3.33810	Bias
15	-0.15122	0.07866	-1.92256	Tidak Bias
16	-0.19261	0.27518	-0.69993	Tidak Bias
17	-0.12758	0.06211	-2.05398	Bias
18	0.09953	0.17614	0.56508	Tidak Bias
19	-0.03988	0.06600	-0.60428	Tidak Bias
20	-0.01745	0.13310	-0.13113	Tidak Bias
21	0.11834	0.11366	1.04121	Tidak Bias
22	-0.01404	0.17939	-0.07828	Tidak Bias
23	-0.10814	0.09622	-1.12388	Tidak Bias
24	-0.15271	0.08345	-1.83002	Tidak Bias
25	-0.00775	0.06655	-0.11643	Tidak Bias
26	-0.08133	0.10485	-0.77568	Tidak Bias
27	-0.02373	0.07763	-0.30567	Tidak Bias
28	0.19626	0.29724	0.66027	Tidak Bias
29	-0.20825	0.16445	-1.26631	Tidak Bias
JUMLAH BIAS BUTIR				5

REPLIKASI KE-15.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
1	R-F	0.03427	R1-R2	0.05439	Tidak Bias
		Non-Uniform	F1-F2	0.04585	
2	R-F	0.02854	R1-R2	0.04338	Tidak Bias
		Non-Uniform	F1-F2	0.00987	
3	R-F	0.03821	R1-R2	0.04471	Tidak Bias
		Non-Uniform	F1-F2	0.04438	
4	R-F	0.02060	R1-R2	0.08028	Tidak Bias
		Uniform	F1-F2	0.04125	
5	R-F	0.04770	R1-R2	0.01242	Bias
		Non-Uniform	F1-F2	0.03887	
6	R-F	0.05238	R1-R2	0.03694	Bias
		Uniform	F1-F2	0.04148	
7	R-F	0.05637	R1-R2	0.04020	Bias
		Non-Uniform	F1-F2	0.02436	
8	R-F	0.05271	R1-R2	0.02026	Bias
		Non-Uniform	F1-F2	0.02518	
9	R-F	0.03280	R1-R2	0.03230	Bias
		Uniform	F1-F2	0.01092	
10	R-F	0.01311	R1-R2	0.05694	Tidak Bias
		Non-Uniform	F1-F2	0.01984	
11	R-F	0.00871	R1-R2	0.02420	Tidak Bias
		Non-Uniform	F1-F2	0.05087	
12	R-F	0.03333	R1-R2	0.03282	Bias
		Uniform	F1-F2	0.03055	
13	R-F	0.05715	R1-R2	0.04151	Bias
		Uniform	F1-F2	0.02867	
14	R-F	0.05520	R1-R2	0.05421	Bias
		Non-Uniform	F1-F2	0.02904	
15	R-F	0.01925	R1-R2	0.05389	Tidak Bias
		Non-Uniform	F1-F2	0.03156	
16	R-F	0.05318	R1-R2	0.06150	Tidak Bias
		Non-Uniform	F1-F2	0.01559	
17	R-F	0.03737	R1-R2	0.02027	Bias
		Uniform	F1-F2	0.03250	
18	R-F	0.00624	R1-R2	0.01244	Tidak Bias
		Non-Uniform	F1-F2	0.03057	
19	R-F	0.01683	R1-R2	0.02354	Tidak Bias
		Non-Uniform	F1-F2	0.05477	
20	R-F	0.04377	R1-R2	0.07020	Tidak Bias
		Non-Uniform	F1-F2	0.02590	
21	R-F	0.06061	R1-R2	0.03711	Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Non-Uniform	F1-F2	0.04168	Tidak Bias
		0.00533	R1-R2	0.05996	
23	R-F	Uniform	F1-F2	0.03235	Tidak Bias
		0.05127	R1-R2	0.06665	
24	R-F	Non-Uniform	F1-F2	0.02995	Tidak Bias
		0.03089	R1-R2	0.03786	
25	R-F	Non-Uniform	F1-F2	0.01717	Tidak Bias
		0.02157	R1-R2	0.02522	
26	R-F	Non-Uniform	F1-F2	0.02961	Tidak Bias
		0.03437	R1-R2	0.05954	
27	R-F	Uniform	F1-F2	0.03799	Tidak Bias
		0.02083	R1-R2	0.03253	
28	R-F	Non-Uniform	F1-F2	0.03353	Tidak Bias
		0.02486	R1-R2	0.03887	
29	R-F	Uniform	F1-F2	0.02038	Tidak Bias
		0.02634	R1-R2	0.03774	
JUMLAH BIAS BUTIR				10	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	3.013517794	7.81472	Tidak Bias
2	6.140441389	7.81472	Tidak Bias
3	2.885526285	7.81472	Tidak Bias
4	3.43476801	7.81472	Tidak Bias
5	5.745485745	7.81472	Tidak Bias
6	8.780814697	7.81472	Bias
7	12.717385	7.81472	Bias
8	9.613107013	7.81472	Bias
9	1.782457537	7.81472	Tidak Bias
10	0.925652977	7.81472	Tidak Bias
11	0.35560757	7.81472	Tidak Bias
12	9.547816775	7.81472	Bias
13	10.9095284	7.81472	Bias
14	12.09925284	7.81472	Bias
15	1.559860518	7.81472	Tidak Bias
16	11.22778367	7.81472	Bias
17	34.9864735	7.81472	Bias
18	0.202759403	7.81472	Tidak Bias
19	0.839963283	7.81472	Tidak Bias
20	7.140015504	7.81472	Tidak Bias
21	6.380252983	7.81472	Tidak Bias
22	0.23219569	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	4.511469368	7.81472	Tidak Bias
24	5.150015293	7.81472	Tidak Bias
25	1.22159947	7.81472	Tidak Bias
26	4.561746094	7.81472	Tidak Bias
27	0.985040825	7.81472	Tidak Bias
28	1.48213323	7.81472	Tidak Bias
29	3.783173617	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			8

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.24477	0.14891	-1.64370	Tidak Bias
2	-0.01837	0.26250	-0.07000	Tidak Bias
3	-0.18873	0.12577	-1.50058	Tidak Bias
4	0.00455	0.16908	0.02689	Tidak Bias
5	0.00252	0.19433	0.01299	Tidak Bias
6	0.12078	0.08533	1.41545	Tidak Bias
7	0.14434	0.10741	1.34390	Tidak Bias
8	0.22103	0.13498	1.63749	Tidak Bias
9	-0.16235	2.41238	-0.06730	Tidak Bias
10	-0.19327	0.30911	-0.62525	Tidak Bias
11	0.08467	0.18116	0.46737	Tidak Bias
12	0.05319	0.07984	0.66613	Tidak Bias
13	-0.01229	0.07054	-0.17426	Tidak Bias
14	-0.22952	0.09638	-2.38138	Bias
15	-0.07913	0.07771	-1.01832	Tidak Bias
16	-0.34737	0.25801	-1.34638	Tidak Bias
17	0.07028	0.05531	1.27074	Tidak Bias
18	0.00694	0.19401	0.03579	Tidak Bias
19	0.05705	0.07173	0.79538	Tidak Bias
20	0.02686	0.15040	0.17858	Tidak Bias
21	0.13660	0.10402	1.31318	Tidak Bias
22	0.03731	0.17252	0.21627	Tidak Bias
23	0.11156	0.09697	1.15046	Tidak Bias
24	-0.03854	0.08631	-0.44658	Tidak Bias
25	0.07984	0.07410	1.07750	Tidak Bias
26	0.06504	0.10754	0.60479	Tidak Bias
27	0.00916	0.07668	0.11947	Tidak Bias
28	0.23012	0.26997	0.85238	Tidak Bias
29	0.05252	0.17161	0.30604	Tidak Bias
JUMLAH BIAS BUTIR				1

REPLIKASI KE-16.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.06037	R1-R2	0.01731	0.06028	Bias
		Non-Uniform	F1-F2	0.06028		
2	R-F	0.03863	R1-R2	0.04176	0.04176	Tidak Bias
		Uniform	F1-F2	0.04008		
3	R-F	0.01828	R1-R2	0.03634	0.03634	Tidak Bias
		Non-Uniform	F1-F2	0.02043		
4	R-F	0.04455	R1-R2	0.05608	0.05608	Tidak Bias
		Non-Uniform	F1-F2	0.03482		
5	R-F	0.02128	R1-R2	0.03184	0.03184	Tidak Bias
		Non-Uniform	F1-F2	0.01600		
6	R-F	0.03608	R1-R2	0.04977	0.04977	Tidak Bias
		Non-Uniform	F1-F2	0.04686		
7	R-F	0.03096	R1-R2	0.00550	0.04093	Tidak Bias
		Non-Uniform	F1-F2	0.04093		
8	R-F	0.04461	R1-R2	0.03492	0.07367	Tidak Bias
		Non-Uniform	F1-F2	0.07367		
9	R-F	0.05622	R1-R2	0.03326	0.03613	Bias
		Uniform	F1-F2	0.03613		
10	R-F	0.01137	R1-R2	0.02603	0.03580	Tidak Bias
		Uniform	F1-F2	0.03580		
11	R-F	0.02985	R1-R2	0.03897	0.03897	Tidak Bias
		Non-Uniform	F1-F2	0.02354		
12	R-F	0.02576	R1-R2	0.03766	0.03766	Tidak Bias
		Non-Uniform	F1-F2	0.02053		
13	R-F	0.04963	R1-R2	0.02747	0.03135	Bias
		Non-Uniform	F1-F2	0.03135		
14	R-F	0.05491	R1-R2	0.03684	0.06572	Tidak Bias
		Non-Uniform	F1-F2	0.06572		
15	R-F	0.06121	R1-R2	0.05803	0.05803	Bias
		Non-Uniform	F1-F2	0.05653		
16	R-F	0.01021	R1-R2	0.05248	0.05248	Tidak Bias
		Non-Uniform	F1-F2	0.03677		
17	R-F	0.02077	R1-R2	0.06419	0.06419	Tidak Bias
		Non-Uniform	F1-F2	0.01436		
18	R-F	0.03435	R1-R2	0.02624	0.02624	Bias
		Uniform	F1-F2	0.01425		
19	R-F	0.02653	R1-R2	0.02418	0.02418	Bias
		Uniform	F1-F2	0.01903		
20	R-F	0.03157	R1-R2	0.05034	0.05200	Tidak Bias
		Uniform	F1-F2	0.05200		
21	R-F	0.03198	R1-R2	0.05233	0.05233	Tidak Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Non-Uniform	F1-F2	0.01274	Bias
		0.04952	R1-R2	0.03271	
23	R-F	Non-Uniform	F1-F2	0.02979	Tidak Bias
		0.03322	R1-R2	0.07469	
24	R-F	Non-Uniform	F1-F2	0.01990	Tidak Bias
		0.01153	R1-R2	0.06530	
25	R-F	Uniform	F1-F2	0.03133	Tidak Bias
		0.01526	R1-R2	0.07118	
26	R-F	Uniform	F1-F2	0.04038	Tidak Bias
		0.02699	R1-R2	0.05183	
27	R-F	Uniform	F1-F2	0.06750	Tidak Bias
		0.01694	R1-R2	0.10553	
28	R-F	Non-Uniform	F1-F2	0.01878	Bias
		0.06148	R1-R2	0.04133	
29	R-F	Uniform	F1-F2	0.02122	Tidak Bias
		0.03023	R1-R2	0.07677	
JUMLAH BIAS BUTIR				8	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	11.18162736	7.81472	Bias
2	7.238502573	7.81472	Tidak Bias
3	1.171625496	7.81472	Tidak Bias
4	8.497725022	7.81472	Bias
5	1.559153461	7.81472	Tidak Bias
6	2.778621771	7.81472	Tidak Bias
7	2.513607972	7.81472	Tidak Bias
8	5.390971514	7.81472	Tidak Bias
9	6.36765468	7.81472	Tidak Bias
10	1.070746271	7.81472	Tidak Bias
11	3.664839082	7.81472	Tidak Bias
12	2.125727672	7.81472	Tidak Bias
13	6.761984101	7.81472	Tidak Bias
14	9.918412211	7.81472	Bias
15	8.389266922	7.81472	Bias
16	0.76680495	7.81472	Tidak Bias
17	0.678307074	7.81472	Tidak Bias
18	7.380818412	7.81472	Tidak Bias
19	3.807520499	7.81472	Tidak Bias
20	3.698207741	7.81472	Tidak Bias
21	2.778036755	7.81472	Tidak Bias
22	6.317779185	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	4.931893498	7.81472	Tidak Bias
24	0.84299765	7.81472	Tidak Bias
25	1.555191445	7.81472	Tidak Bias
26	1.973813927	7.81472	Tidak Bias
27	0.804297607	7.81472	Tidak Bias
28	11.42379857	7.81472	Bias
29	2.729696238	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			5

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.41819	0.14420	-2.90016	Bias
2	0.26677	0.23965	1.11316	Tidak Bias
3	-0.11421	0.11687	-0.97726	Tidak Bias
4	0.19966	0.17916	1.11442	Tidak Bias
5	0.16507	0.18485	0.89298	Tidak Bias
6	0.11847	0.08002	1.48048	Tidak Bias
7	0.12728	0.11152	1.14131	Tidak Bias
8	-0.39022	0.17014	-2.29359	Bias
9	0.40948	2.07909	0.19695	Tidak Bias
10	-0.04961	0.34817	-0.14248	Tidak Bias
11	0.20649	0.17785	1.16105	Tidak Bias
12	0.10950	0.08388	1.30543	Tidak Bias
13	-0.13584	0.07431	-1.82800	Tidak Bias
14	-0.19666	0.09028	-2.17835	Bias
15	-0.20140	0.08375	-2.40473	Bias
16	-0.09598	0.26365	-0.36403	Tidak Bias
17	-0.01214	0.05741	-0.21151	Tidak Bias
18	0.16953	0.16510	1.02685	Tidak Bias
19	-0.04908	0.07491	-0.65511	Tidak Bias
20	0.07669	0.12746	0.60169	Tidak Bias
21	0.12877	0.12405	1.03810	Tidak Bias
22	-0.25211	0.18766	-1.34340	Tidak Bias
23	-0.02615	0.09730	-0.26876	Tidak Bias
24	0.00764	0.07892	0.09680	Tidak Bias
25	-0.04582	0.07497	-0.61116	Tidak Bias
26	-0.01827	0.10792	-0.16931	Tidak Bias
27	-0.03118	0.08317	-0.37486	Tidak Bias
28	0.36757	0.24246	1.51603	Tidak Bias
29	-0.31606	0.20186	-1.56569	Tidak Bias
JUMLAH BIAS BUTIR				4

REPLIKASI KE-17.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok		Indeks Subkelompok		Maksimum	Kesimpulan
1	R-F	0.03735	R1-R2	0.01852	0.01852	Bias
		Non-Uniform	F1-F2	0.01592		
2	R-F	0.04196	R1-R2	0.03597	0.03597	Bias
		Non-Uniform	F1-F2	0.01906		
3	R-F	0.05590	R1-R2	0.02797	0.06990	Tidak Bias
		Non-Uniform	F1-F2	0.06990		
4	R-F	0.04455	R1-R2	0.01795	0.04923	Tidak Bias
		Non-Uniform	F1-F2	0.04923		
5	R-F	0.01449	R1-R2	0.05028	0.05376	Tidak Bias
		Non-Uniform	F1-F2	0.05376		
6	R-F	0.02196	R1-R2	0.04301	0.04301	Tidak Bias
		Non-Uniform	F1-F2	0.00973		
7	R-F	0.04226	R1-R2	0.01976	0.01976	Bias
		Non-Uniform	F1-F2	0.01190		
8	R-F	0.03457	R1-R2	0.02855	0.06669	Tidak Bias
		Non-Uniform	F1-F2	0.06669		
9	R-F	0.04821	R1-R2	0.03046	0.03046	Bias
		Uniform	F1-F2	0.01650		
10	R-F	0.01292	R1-R2	0.00983	0.02239	Tidak Bias
		Non-Uniform	F1-F2	0.02239		
11	R-F	0.01133	R1-R2	0.01271	0.04118	Tidak Bias
		Non-Uniform	F1-F2	0.04118		
12	R-F	0.04348	R1-R2	0.02473	0.02473	Bias
		Non-Uniform	F1-F2	0.01771		
13	R-F	0.05502	R1-R2	0.01183	0.03961	Bias
		Non-Uniform	F1-F2	0.03961		
14	R-F	0.02470	R1-R2	0.03756	0.05716	Tidak Bias
		Non-Uniform	F1-F2	0.05716		
15	R-F	0.02861	R1-R2	0.04979	0.04979	Tidak Bias
		Uniform	F1-F2	0.02730		
16	R-F	0.02984	R1-R2	0.00550	0.03656	Tidak Bias
		Non-Uniform	F1-F2	0.03656		
17	R-F	0.01771	R1-R2	0.02271	0.02271	Tidak Bias
		Non-Uniform	F1-F2	0.01948		
18	R-F	0.03838	R1-R2	0.03497	0.03497	Bias
		Non-Uniform	F1-F2	0.02892		
19	R-F	0.02347	R1-R2	0.01983	0.04183	Tidak Bias
		Non-Uniform	F1-F2	0.04183		
20	R-F	0.03146	R1-R2	0.04748	0.04748	Tidak Bias
		Non-Uniform	F1-F2	0.03212		
21	R-F	0.04430	R1-R2	0.05872	0.05872	Tidak Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Non-Uniform	F1-F2	0.04501	Bias
		0.03357	R1-R2	0.02808	
23	R-F	Non-Uniform	F1-F2	0.01500	Bias
		0.03745	R1-R2	0.02359	
24	R-F	Non-Uniform	F1-F2	0.02957	Tidak Bias
		0.01665	R1-R2	0.03211	
25	R-F	Non-Uniform	F1-F2	0.01580	Bias
		0.02758	R1-R2	0.01808	
26	R-F	Non-Uniform	F1-F2	0.01513	Tidak Bias
		0.02240	R1-R2	0.02268	
27	R-F	Non-Uniform	F1-F2	0.04269	Tidak Bias
		0.02968	R1-R2	0.04375	
28	R-F	Non-Uniform	F1-F2	0.06267	Bias
		0.03871	R1-R2	0.01002	
29	R-F	Uniform	F1-F2	0.00408	Tidak Bias
		0.02947	R1-R2	0.04144	
		Non-Uniform	F1-F2	0.02551	
JUMLAH BIAS BUTIR				11	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	4.231695174	7.81472	Tidak Bias
2	11.28343277	7.81472	Bias
3	6.092744416	7.81472	Tidak Bias
4	10.41610357	7.81472	Bias
5	0.439864337	7.81472	Tidak Bias
6	1.501070679	7.81472	Tidak Bias
7	6.860011116	7.81472	Tidak Bias
8	5.083251029	7.81472	Tidak Bias
9	2.699105833	7.81472	Tidak Bias
10	0.381967453	7.81472	Tidak Bias
11	0.297670348	7.81472	Tidak Bias
12	4.644263076	7.81472	Tidak Bias
13	5.22401225	7.81472	Tidak Bias
14	1.594280113	7.81472	Tidak Bias
15	2.702470335	7.81472	Tidak Bias
16	1.691802456	7.81472	Tidak Bias
17	2.084369812	7.81472	Tidak Bias
18	9.464625318	7.81472	Bias
19	1.496672827	7.81472	Tidak Bias
20	2.86887584	7.81472	Tidak Bias
21	3.444401022	7.81472	Tidak Bias
22	3.676982669	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	3.057958497	7.81472	Tidak Bias
24	0.936273306	7.81472	Tidak Bias
25	3.007679449	7.81472	Tidak Bias
26	1.382335003	7.81472	Tidak Bias
27	2.267714036	7.81472	Tidak Bias
28	4.063055476	7.81472	Tidak Bias
29	4.113038512	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			3

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.20002	0.14305	-1.39820	Tidak Bias
2	0.06150	0.21904	0.28076	Tidak Bias
3	-0.21724	0.11165	-1.94565	Tidak Bias
4	0.11567	0.15696	0.73688	Tidak Bias
5	-0.01906	0.18365	-0.10380	Tidak Bias
6	0.08338	0.08745	0.95350	Tidak Bias
7	0.11592	0.11888	0.97506	Tidak Bias
8	0.06654	0.15430	0.43124	Tidak Bias
9	-0.10283	1.43115	-0.07185	Tidak Bias
10	-0.16178	0.32637	-0.49570	Tidak Bias
11	-0.05840	0.16382	-0.35647	Tidak Bias
12	0.18367	0.08695	2.11224	Bias
13	0.05642	0.07575	0.74480	Tidak Bias
14	-0.13072	0.10967	-1.19203	Tidak Bias
15	0.00249	0.07338	0.03398	Tidak Bias
16	-0.21939	0.23846	-0.92006	Tidak Bias
17	0.05636	0.05960	0.94572	Tidak Bias
18	0.13554	0.17455	0.77651	Tidak Bias
19	-0.01315	0.07457	-0.17630	Tidak Bias
20	0.03181	0.14031	0.22670	Tidak Bias
21	0.12211	0.09994	1.22189	Tidak Bias
22	-0.15741	0.14307	-1.10023	Tidak Bias
23	0.08572	0.09697	0.88397	Tidak Bias
24	-0.06330	0.08573	-0.73840	Tidak Bias
25	0.04891	0.07543	0.64843	Tidak Bias
26	0.06033	0.10670	0.56536	Tidak Bias
27	0.00175	0.08357	0.02098	Tidak Bias
28	0.19479	0.24817	0.78494	Tidak Bias
29	-0.07962	0.18506	-0.43022	Tidak Bias
JUMLAH BIAS BUTIR				1

REPLIKASI KE-18.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.05413	R1-R2	0.05445	0.05445	Tidak Bias
		Non-Uniform	F1-F2	0.02096		
2	R-F	0.02335	R1-R2	0.02131	0.02131	Bias
		Uniform	F1-F2	0.01841		
3	R-F	0.02996	R1-R2	0.03143	0.03143	Tidak Bias
		Non-Uniform	F1-F2	0.02371		
4	R-F	0.02053	R1-R2	0.02958	0.02958	Tidak Bias
		Uniform	F1-F2	0.01092		
5	R-F	0.04599	R1-R2	0.03661	0.03883	Bias
		Non-Uniform	F1-F2	0.03883		
6	R-F	0.01481	R1-R2	0.03127	0.05479	Tidak Bias
		Uniform	F1-F2	0.05479		
7	R-F	0.04457	R1-R2	0.01835	0.05296	Tidak Bias
		Non-Uniform	F1-F2	0.05296		
8	R-F	0.05059	R1-R2	0.05856	0.05856	Tidak Bias
		Non-Uniform	F1-F2	0.01645		
9	R-F	0.06245	R1-R2	0.04976	0.04976	Bias
		Uniform	F1-F2	0.01303		
10	R-F	0.01512	R1-R2	0.06412	0.06412	Tidak Bias
		Uniform	F1-F2	0.03303		
11	R-F	0.01129	R1-R2	0.02949	0.02949	Tidak Bias
		Uniform	F1-F2	0.02663		
12	R-F	0.02451	R1-R2	0.01767	0.05270	Tidak Bias
		Non-Uniform	F1-F2	0.05270		
13	R-F	0.03584	R1-R2	0.04462	0.04462	Tidak Bias
		Non-Uniform	F1-F2	0.03026		
14	R-F	0.04395	R1-R2	0.05236	0.05841	Tidak Bias
		Non-Uniform	F1-F2	0.05841		
15	R-F	0.03403	R1-R2	0.05957	0.05957	Tidak Bias
		Uniform	F1-F2	0.04597		
16	R-F	0.01538	R1-R2	0.04369	0.04369	Tidak Bias
		Non-Uniform	F1-F2	0.03194		
17	R-F	0.01723	R1-R2	0.05362	0.05362	Tidak Bias
		Non-Uniform	F1-F2	0.02907		
18	R-F	0.02140	R1-R2	0.00696	0.04310	Tidak Bias
		Non-Uniform	F1-F2	0.04310		
19	R-F	0.02150	R1-R2	0.03860	0.03860	Tidak Bias
		Uniform	F1-F2	0.02642		
20	R-F	0.02172	R1-R2	0.01746	0.04542	Tidak Bias
		Non-Uniform	F1-F2	0.04542		
21	R-F	0.02819	R1-R2	0.05256	0.06518	Tidak Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Non-Uniform	F1-F2	0.06518	Tidak Bias
		0.03685	R1-R2	0.01986	
23	R-F	Non-Uniform	F1-F2	0.05918	Tidak Bias
		0.04841	R1-R2	0.05146	
24	R-F	Non-Uniform	F1-F2	0.03586	Tidak Bias
		0.02116	R1-R2	0.02785	
25	R-F	Non-Uniform	F1-F2	0.01529	Tidak Bias
		0.05187	R1-R2	0.03809	
26	R-F	Non-Uniform	F1-F2	0.04402	Tidak Bias
		0.05385	R1-R2	0.05604	
27	R-F	Non-Uniform	F1-F2	0.04992	Tidak Bias
		0.01784	R1-R2	0.03886	
28	R-F	Uniform	F1-F2	0.05624	Tidak Bias
		0.02549	R1-R2	0.03463	
29	R-F	Uniform	F1-F2	0.02691	Tidak Bias
		0.03781	R1-R2	0.03705	
		Non-Uniform	F1-F2	0.02226	Bias
JUMLAH BIAS BUTIR				5	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	7.694179466	7.81472	Tidak Bias
2	2.755626241	7.81472	Tidak Bias
3	2.903225799	7.81472	Tidak Bias
4	1.639859325	7.81472	Tidak Bias
5	5.510356914	7.81472	Tidak Bias
6	1.225314412	7.81472	Tidak Bias
7	9.063490388	7.81472	Bias
8	7.790721801	7.81472	Tidak Bias
9	12.61032411	7.81472	Bias
10	1.942111387	7.81472	Tidak Bias
11	1.514217156	7.81472	Tidak Bias
12	2.602387642	7.81472	Tidak Bias
13	5.99209711	7.81472	Tidak Bias
14	4.283779083	7.81472	Tidak Bias
15	3.194254241	7.81472	Tidak Bias
16	0.986843317	7.81472	Tidak Bias
17	1.142203316	7.81472	Tidak Bias
18	1.411669553	7.81472	Tidak Bias
19	2.926939228	7.81472	Tidak Bias
20	1.731166999	7.81472	Tidak Bias
21	3.058986241	7.81472	Tidak Bias
22	4.646348065	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	5.072109292	7.81472	Tidak Bias
24	1.481635656	7.81472	Tidak Bias
25	9.362717332	7.81472	Bias
26	7.146584899	7.81472	Tidak Bias
27	1.063776171	7.81472	Tidak Bias
28	1.717687621	7.81472	Tidak Bias
29	5.493212735	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			3

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.37098	0.13756	-2.69687	Bias
2	0.17281	0.23649	0.73072	Tidak Bias
3	-0.12324	0.12007	-1.02634	Tidak Bias
4	0.14143	0.18567	0.76175	Tidak Bias
5	-0.02793	0.19081	-0.14638	Tidak Bias
6	0.01588	0.08179	0.19420	Tidak Bias
7	0.08363	0.11724	0.71328	Tidak Bias
8	-0.17537	0.17379	-1.00907	Tidak Bias
9	0.09832	0.94414	0.10413	Tidak Bias
10	-0.08131	0.31689	-0.25658	Tidak Bias
11	0.07998	0.17472	0.45777	Tidak Bias
12	0.08799	0.08728	1.00822	Tidak Bias
13	-0.11676	0.07724	-1.51165	Tidak Bias
14	0.12975	0.10134	1.28028	Tidak Bias
15	-0.07044	0.07732	-0.91104	Tidak Bias
16	-0.16826	0.28304	-0.59447	Tidak Bias
17	-0.01010	0.05924	-0.17045	Tidak Bias
18	0.12212	0.19209	0.63574	Tidak Bias
19	-0.04839	0.07722	-0.62659	Tidak Bias
20	0.05848	0.13998	0.41777	Tidak Bias
21	0.05287	0.11797	0.44821	Tidak Bias
22	0.16813	0.14783	1.13734	Tidak Bias
23	-0.18098	0.08977	-2.01618	Bias
24	0.02314	0.08921	0.25939	Tidak Bias
25	0.04196	0.07382	0.56844	Tidak Bias
26	0.27303	0.10319	2.64596	Bias
27	-0.05532	0.08431	-0.65622	Tidak Bias
28	0.10432	0.29582	0.35264	Tidak Bias
29	-0.22478	0.17563	-1.27987	Tidak Bias
JUMLAH BIAS BUTIR				3

REPLIKASI KE-19.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.03967	R1-R2	0.02703	0.02703	Bias
		Non-Uniform	F1-F2	0.02587		
2	R-F	0.06066	R1-R2	0.01779	0.01779	Bias
		Non-Uniform	F1-F2	0.00903		
3	R-F	0.02737	R1-R2	0.07218	0.07218	Tidak Bias
		Non-Uniform	F1-F2	0.01552		
4	R-F	0.02638	R1-R2	0.01209	0.02214	Bias
		Non-Uniform	F1-F2	0.02214		
5	R-F	0.03796	R1-R2	0.03834	0.06645	Tidak Bias
		Non-Uniform	F1-F2	0.06645		
6	R-F	0.03728	R1-R2	0.01582	0.09421	Tidak Bias
		Uniform	F1-F2	0.09421		
7	R-F	0.04937	R1-R2	0.00346	0.02457	Bias
		Non-Uniform	F1-F2	0.02457		
8	R-F	0.06217	R1-R2	0.02394	0.04421	Bias
		Non-Uniform	F1-F2	0.04421		
9	R-F	0.07685	R1-R2	0.03725	0.03725	Bias
		Non-Uniform	F1-F2	0.00363		
10	R-F	0.00922	R1-R2	0.05724	0.05724	Tidak Bias
		Uniform	F1-F2	0.03214		
11	R-F	0.00301	R1-R2	0.01981	0.05095	Tidak Bias
		Uniform	F1-F2	0.05095		
12	R-F	0.05478	R1-R2	0.04320	0.04320	Bias
		Uniform	F1-F2	0.03564		
13	R-F	0.03060	R1-R2	0.04575	0.04575	Tidak Bias
		Non-Uniform	F1-F2	0.02519		
14	R-F	0.05683	R1-R2	0.01391	0.04044	Bias
		Uniform	F1-F2	0.04044		
15	R-F	0.03505	R1-R2	0.02366	0.02366	Bias
		Uniform	F1-F2	0.02177		
16	R-F	0.03583	R1-R2	0.03799	0.03799	Tidak Bias
		Non-Uniform	F1-F2	0.02573		
17	R-F	0.05753	R1-R2	0.02631	0.04665	Bias
		Uniform	F1-F2	0.04665		
18	R-F	0.01041	R1-R2	0.02237	0.02237	Tidak Bias
		Non-Uniform	F1-F2	0.01386		
19	R-F	0.02031	R1-R2	0.02769	0.02769	Tidak Bias
		Uniform	F1-F2	0.02704		
20	R-F	0.03789	R1-R2	0.04308	0.04308	Tidak Bias
		Non-Uniform	F1-F2	0.02983		
21	R-F	0.05861	R1-R2	0.03505	0.04521	Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
22	R-F	Uniform	F1-F2	0.04521	0.03815	Tidak Bias
		0.03348	R1-R2	0.03039		
23	R-F	Non-Uniform	F1-F2	0.03815	0.06734	Tidak Bias
		0.02006	R1-R2	0.06734		
24	R-F	Non-Uniform	F1-F2	0.03532	0.05312	Tidak Bias
		0.03070	R1-R2	0.05312		
25	R-F	Non-Uniform	F1-F2	0.03036	0.04036	Bias
		0.06948	R1-R2	0.04036		
26	R-F	Non-Uniform	F1-F2	0.02957	0.03749	Bias
		0.04856	R1-R2	0.02176		
27	R-F	Non-Uniform	F1-F2	0.03749	0.02880	Tidak Bias
		0.01741	R1-R2	0.02880		
28	R-F	Non-Uniform	F1-F2	0.01833	0.03899	Tidak Bias
		0.02387	R1-R2	0.01841		
29	R-F	Uniform	F1-F2	0.03899	0.05278	Tidak Bias
		0.03731	R1-R2	0.02495		
Jumlah Bias Butir					13	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	3.702308751	7.81472	Tidak Bias
2	27.63116713	7.81472	Bias
3	6.865123919	7.81472	Tidak Bias
4	3.526847173	7.81472	Tidak Bias
5	5.277963145	7.81472	Tidak Bias
6	4.879354302	7.81472	Tidak Bias
7	11.44898988	7.81472	Bias
8	10.92396375	7.81472	Bias
9	5.233058957	7.81472	Tidak Bias
10	0.394311812	7.81472	Tidak Bias
11	0.529615424	7.81472	Tidak Bias
12	19.53025926	7.81472	Bias
13	1.631096922	7.81472	Tidak Bias
14	7.839288139	7.81472	Bias
15	4.582900133	7.81472	Tidak Bias
16	2.271449932	7.81472	Tidak Bias
17	35.17300435	7.81472	Bias
18	1.663080653	7.81472	Tidak Bias
19	2.223680174	7.81472	Tidak Bias
20	6.122565473	7.81472	Tidak Bias
21	18.37235352	7.81472	Bias
22	3.425294313	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	2.28043429	7.81472	Tidak Bias
24	3.059520801	7.81472	Tidak Bias
25	17.61923532	7.81472	Bias
26	7.056821283	7.81472	Tidak Bias
27	0.911068906	7.81472	Tidak Bias
28	1.190144736	7.81472	Tidak Bias
29	9.21333815	7.81472	Bias
JUMLAH BIAS BUTIR			9

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.23968	0.14137	-1.69534	Tidak Bias
2	-0.20740	0.24057	-0.86209	Tidak Bias
3	-0.04272	0.12590	-0.33932	Tidak Bias
4	0.02339	0.17952	0.13032	Tidak Bias
5	0.09622	0.17247	0.55790	Tidak Bias
6	0.12825	0.08036	1.59599	Tidak Bias
7	0.10329	0.11139	0.92728	Tidak Bias
8	-0.21035	0.18406	-1.14281	Tidak Bias
9	-0.36429	1.43696	-0.25352	Tidak Bias
10	-0.15123	0.31937	-0.47352	Tidak Bias
11	-0.01697	0.16623	-0.10207	Tidak Bias
12	0.14881	0.07656	1.94358	Tidak Bias
13	0.06652	0.06619	1.00491	Tidak Bias
14	0.09321	0.10006	0.93151	Tidak Bias
15	0.07980	0.07926	1.00680	Tidak Bias
16	-0.19389	0.24558	-0.78952	Tidak Bias
17	0.10792	0.05566	1.93904	Tidak Bias
18	-0.11171	0.18982	-0.58850	Tidak Bias
19	0.07800	0.06936	1.12448	Tidak Bias
20	0.01238	0.13848	0.08941	Tidak Bias
21	0.22990	0.10780	2.13263	Bias
22	-0.10351	0.15230	-0.67966	Tidak Bias
23	0.01423	0.10023	0.14199	Tidak Bias
24	-0.01101	0.07773	-0.14160	Tidak Bias
25	0.15523	0.07673	2.02297	Bias
26	0.18869	0.10621	1.77653	Tidak Bias
27	-0.02001	0.08298	-0.24117	Tidak Bias
28	0.12688	0.26613	0.47676	Tidak Bias
29	0.02004	0.16908	0.11851	Tidak Bias
JUMLAH BIAS BUTIR				2

REPLIKASI KE-20.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
1	R-F	0.04222	R1-R2	0.02829	Bias
		Non-Uniform	F1-F2	0.01881	
2	R-F	0.04567	R1-R2	0.03815	Bias
		Non-Uniform	F1-F2	0.02748	
3	R-F	0.05270	R1-R2	0.04011	Tidak Bias
		Non-Uniform	F1-F2	0.06938	
4	R-F	0.03811	R1-R2	0.08835	Tidak Bias
		Non-Uniform	F1-F2	0.05189	
5	R-F	0.01732	R1-R2	0.06840	Tidak Bias
		Non-Uniform	F1-F2	0.00958	
6	R-F	0.03334	R1-R2	0.06476	Tidak Bias
		Non-Uniform	F1-F2	0.02623	
7	R-F	0.03840	R1-R2	0.02903	Tidak Bias
		Uniform	F1-F2	0.04376	
8	R-F	0.03532	R1-R2	0.04954	Tidak Bias
		Non-Uniform	F1-F2	0.01222	
9	R-F	0.05291	R1-R2	0.05263	Bias
		Non-Uniform	F1-F2	0.00764	
10	R-F	0.01619	R1-R2	0.02942	Tidak Bias
		Non-Uniform	F1-F2	0.01552	
11	R-F	0.02449	R1-R2	0.05480	Tidak Bias
		Non-Uniform	F1-F2	0.03388	
12	R-F	0.03466	R1-R2	0.03297	Bias
		Non-Uniform	F1-F2	0.02052	
13	R-F	0.04451	R1-R2	0.04813	Tidak Bias
		Uniform	F1-F2	0.07748	
14	R-F	0.02464	R1-R2	0.05869	Tidak Bias
		Non-Uniform	F1-F2	0.07317	
15	R-F	0.04264	R1-R2	0.04923	Tidak Bias
		Non-Uniform	F1-F2	0.00687	
16	R-F	0.00533	R1-R2	0.04164	Tidak Bias
		Non-Uniform	F1-F2	0.03605	
17	R-F	0.04100	R1-R2	0.09098	Tidak Bias
		Non-Uniform	F1-F2	0.04208	
18	R-F	0.07365	R1-R2	0.02900	Bias
		Non-Uniform	F1-F2	0.02518	
19	R-F	0.01379	R1-R2	0.07820	Tidak Bias
		Non-Uniform	F1-F2	0.02197	
20	R-F	0.03757	R1-R2	0.02524	Bias
		Non-Uniform	F1-F2	0.03000	
21	R-F	0.02867	R1-R2	0.06892	Tidak Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Non-Uniform	F1-F2	0.05429	Tidak Bias
		0.01754	R1-R2	0.05971	
23	R-F	Non-Uniform	F1-F2	0.08799	Tidak Bias
		0.02239	R1-R2	0.04052	
24	R-F	Non-Uniform	F1-F2	0.03835	Tidak Bias
		0.02553	R1-R2	0.06097	
25	R-F	Non-Uniform	F1-F2	0.04361	Tidak Bias
		0.02040	R1-R2	0.05590	
26	R-F	Uniform	F1-F2	0.03915	Tidak Bias
		0.02676	R1-R2	0.04061	
27	R-F	Non-Uniform	F1-F2	0.04370	Tidak Bias
		0.01877	R1-R2	0.03131	
28	R-F	Uniform	F1-F2	0.03446	Tidak Bias
		0.04455	R1-R2	0.07629	
29	R-F	Uniform	F1-F2	0.02701	Bias
		0.05610	R1-R2	0.05270	
JUMLAH BIAS BUTIR				7	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	5.012133296	7.81472	Tidak Bias
2	15.54460114	7.81472	Bias
3	5.624953131	7.81472	Tidak Bias
4	7.68957944	7.81472	Tidak Bias
5	0.975105829	7.81472	Tidak Bias
6	5.268382821	7.81472	Tidak Bias
7	6.734507087	7.81472	Tidak Bias
8	5.107376613	7.81472	Tidak Bias
9	2.107163605	7.81472	Tidak Bias
10	3.543234035	7.81472	Tidak Bias
11	1.996453844	7.81472	Tidak Bias
12	5.58032508	7.81472	Tidak Bias
13	4.192736625	7.81472	Tidak Bias
14	1.723102829	7.81472	Tidak Bias
15	4.404485223	7.81472	Tidak Bias
16	0.514400366	7.81472	Tidak Bias
17	4.282348072	7.81472	Tidak Bias
18	12.19715576	7.81472	Bias
19	0.91252504	7.81472	Tidak Bias
20	5.957137055	7.81472	Tidak Bias
21	1.42092644	7.81472	Tidak Bias
22	1.523665261	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	2.512801147	7.81472	Tidak Bias
24	3.622366289	7.81472	Tidak Bias
25	1.713430846	7.81472	Tidak Bias
26	1.74966585	7.81472	Tidak Bias
27	0.852071622	7.81472	Tidak Bias
28	5.347039107	7.81472	Tidak Bias
29	8.540276452	7.81472	Bias
JUMLAH BIAS BUTIR			3

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.27325	0.13653	-2.00137	Bias
2	-0.18290	0.25221	-0.72517	Tidak Bias
3	-0.20304	0.12346	-1.64449	Tidak Bias
4	0.03511	0.16187	0.21687	Tidak Bias
5	0.07162	0.16874	0.42445	Tidak Bias
6	0.08650	0.07376	1.17283	Tidak Bias
7	0.11263	0.10515	1.07110	Tidak Bias
8	-0.07235	0.14072	-0.51411	Tidak Bias
9	-0.08043	1.64903	-0.04877	Tidak Bias
10	-0.21823	0.34827	-0.62659	Tidak Bias
11	0.12383	0.18451	0.67113	Tidak Bias
12	0.11068	0.07877	1.40500	Tidak Bias
13	0.03329	0.07031	0.47345	Tidak Bias
14	0.00302	0.10804	0.02795	Tidak Bias
15	-0.05558	0.07290	-0.76233	Tidak Bias
16	-0.10568	0.25692	-0.41133	Tidak Bias
17	0.05758	0.05588	1.03052	Tidak Bias
18	0.27662	0.17799	1.55412	Tidak Bias
19	-0.01770	0.07254	-0.24405	Tidak Bias
20	0.01584	0.13499	0.11734	Tidak Bias
21	0.05909	0.10732	0.55056	Tidak Bias
22	0.03990	0.14637	0.27259	Tidak Bias
23	-0.08133	0.08931	-0.91061	Tidak Bias
24	0.03795	0.07991	0.47485	Tidak Bias
25	0.02272	0.06553	0.34677	Tidak Bias
26	0.11558	0.09842	1.17433	Tidak Bias
27	-0.00025	0.07441	-0.00335	Tidak Bias
28	0.25172	0.23059	1.09163	Tidak Bias
29	-0.16295	0.17435	-0.93461	Tidak Bias
JUMLAH BIAS BUTIR				1

REPLIKASI KE-21.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.04504	R1-R2	0.05228	0.06228	Tidak Bias
		Non-Uniform	F1-F2	0.06228		
2	R-F	0.03869	R1-R2	0.02760	0.02760	Bias
		Non-Uniform	F1-F2	0.02463		
3	R-F	0.04906	R1-R2	0.03830	0.03830	Bias
		Non-Uniform	F1-F2	0.03256		
4	R-F	0.02680	R1-R2	0.05830	0.05830	Tidak Bias
		Non-Uniform	F1-F2	0.03514		
5	R-F	0.03496	R1-R2	0.09570	0.09570	Tidak Bias
		Non-Uniform	F1-F2	0.03288		
6	R-F	0.07236	R1-R2	0.08087	0.08087	Tidak Bias
		Non-Uniform	F1-F2	0.06377		
7	R-F	0.04608	R1-R2	0.06174	0.06174	Tidak Bias
		Non-Uniform	F1-F2	0.02324		
8	R-F	0.02879	R1-R2	0.03232	0.06582	Tidak Bias
		Non-Uniform	F1-F2	0.06582		
9	R-F	0.04192	R1-R2	0.05742	0.05742	Tidak Bias
		Uniform	F1-F2	0.03192		
10	R-F	0.02554	R1-R2	0.03196	0.03196	Tidak Bias
		Non-Uniform	F1-F2	0.00781		
11	R-F	0.03710	R1-R2	0.01848	0.03618	Bias
		Non-Uniform	F1-F2	0.03618		
12	R-F	0.03077	R1-R2	0.07584	0.07584	Tidak Bias
		Uniform	F1-F2	0.00679		
13	R-F	0.05869	R1-R2	0.04063	0.04063	Bias
		Uniform	F1-F2	0.02851		
14	R-F	0.02663	R1-R2	0.07182	0.07182	Tidak Bias
		Non-Uniform	F1-F2	0.02312		
15	R-F	0.05020	R1-R2	0.07376	0.07376	Tidak Bias
		Non-Uniform	F1-F2	0.02109		
16	R-F	0.01774	R1-R2	0.04882	0.04882	Tidak Bias
		Non-Uniform	F1-F2	0.02040		
17	R-F	0.01968	R1-R2	0.10759	0.10759	Tidak Bias
		Non-Uniform	F1-F2	0.04555		
18	R-F	0.01809	R1-R2	0.04363	0.04363	Tidak Bias
		Uniform	F1-F2	0.03645		
19	R-F	0.01458	R1-R2	0.06787	0.06787	Tidak Bias
		Non-Uniform	F1-F2	0.01162		
20	R-F	0.02372	R1-R2	0.03526	0.06426	Tidak Bias
		Non-Uniform	F1-F2	0.06426		
21	R-F	0.03129	R1-R2	0.05668	0.05668	Tidak Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Non-Uniform	F1-F2	0.02211	Tidak Bias
		0.02403	R1-R2	0.04321	
23	R-F	Non-Uniform	F1-F2	0.02662	Tidak Bias
		0.01309	R1-R2	0.01146	
24	R-F	Non-Uniform	F1-F2	0.08419	Tidak Bias
		0.01330	R1-R2	0.06919	
25	R-F	Non-Uniform	F1-F2	0.02186	Tidak Bias
		0.02196	R1-R2	0.05720	
26	R-F	Non-Uniform	F1-F2	0.02895	Tidak Bias
		0.02294	R1-R2	0.04159	
27	R-F	Non-Uniform	F1-F2	0.02537	Bias
		0.03489	R1-R2	0.00484	
28	R-F	Non-Uniform	F1-F2	0.02287	Bias
		0.03364	R1-R2	0.02379	
29	R-F	Uniform	F1-F2	0.01629	Bias
		0.04508	R1-R2	0.02212	
		Non-Uniform	F1-F2	0.03512	
JUMLAH BIAS BUTIR				7	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	8.007208403	7.81472	Bias
2	8.855200266	7.81472	Bias
3	6.598989039	7.81472	Tidak Bias
4	2.268905337	7.81472	Tidak Bias
5	3.713728831	7.81472	Tidak Bias
6	8.989104878	7.81472	Bias
7	8.223010711	7.81472	Bias
8	2.494146223	7.81472	Tidak Bias
9	3.276077987	7.81472	Tidak Bias
10	6.090161962	7.81472	Tidak Bias
11	10.10317343	7.81472	Bias
12	4.537954028	7.81472	Tidak Bias
13	19.9261632	7.81472	Bias
14	2.772063211	7.81472	Tidak Bias
15	7.366257825	7.81472	Tidak Bias
16	1.549652058	7.81472	Tidak Bias
17	2.995562857	7.81472	Tidak Bias
18	1.738681317	7.81472	Tidak Bias
19	0.968453846	7.81472	Tidak Bias
20	2.14144188	7.81472	Tidak Bias
21	1.716805569	7.81472	Tidak Bias
22	1.181336622	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	0.488918051	7.81472	Tidak Bias
24	0.60904259	7.81472	Tidak Bias
25	1.985891657	7.81472	Tidak Bias
26	1.554605726	7.81472	Tidak Bias
27	2.829888522	7.81472	Tidak Bias
28	2.674260878	7.81472	Tidak Bias
29	8.950512837	7.81472	Bias
JUMLAH BIAS BUTIR			7

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.32472	0.13264	-2.44819	Bias
2	0.05631	0.27521	0.20460	Tidak Bias
3	-0.17550	0.13293	-1.32023	Tidak Bias
4	0.18069	0.18016	1.00291	Tidak Bias
5	0.08627	0.17067	0.50549	Tidak Bias
6	0.17455	0.08410	2.07546	Bias
7	0.12172	0.10590	1.14934	Tidak Bias
8	-0.22961	0.18690	-1.22854	Tidak Bias
9	-0.16918	2.12402	-0.07965	Tidak Bias
10	-0.60503	0.38173	-1.58498	Tidak Bias
11	0.26792	0.18897	1.41778	Tidak Bias
12	0.10750	0.08894	1.20867	Tidak Bias
13	-0.03473	0.07320	-0.47447	Tidak Bias
14	-0.14473	0.10619	-1.36284	Tidak Bias
15	0.05086	0.08129	0.62564	Tidak Bias
16	-0.20799	0.27754	-0.74940	Tidak Bias
17	-0.03730	0.06071	-0.61436	Tidak Bias
18	0.09561	0.17000	0.56243	Tidak Bias
19	0.02279	0.07382	0.30875	Tidak Bias
20	0.01933	0.13519	0.14301	Tidak Bias
21	0.09936	0.12102	0.82101	Tidak Bias
22	0.12501	0.17100	0.73101	Tidak Bias
23	0.04816	0.09486	0.50770	Tidak Bias
24	0.04459	0.08460	0.52714	Tidak Bias
25	0.05083	0.07534	0.67474	Tidak Bias
26	0.10553	0.11525	0.91563	Tidak Bias
27	-0.00023	0.08457	-0.00269	Tidak Bias
28	0.24506	0.31211	0.78516	Tidak Bias
29	0.02694	0.17539	0.15358	Tidak Bias
JUMLAH BIAS BUTIR				2

REPLIKASI KE-22.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.04215	R1-R2	0.01536	0.05197	Tidak Bias
		Non-Uniform	F1-F2	0.05197		
2	R-F	0.03573	R1-R2	0.01790	0.01790	Bias
		Uniform	F1-F2	0.00642		
3	R-F	0.01820	R1-R2	0.01877	0.01925	Tidak Bias
		Non-Uniform	F1-F2	0.01925		
4	R-F	0.00755	R1-R2	0.02191	0.02191	Tidak Bias
		Non-Uniform	F1-F2	0.00870		
5	R-F	0.04702	R1-R2	0.04364	0.05402	Tidak Bias
		Non-Uniform	F1-F2	0.05402		
6	R-F	0.00887	R1-R2	0.04280	0.04280	Tidak Bias
		Non-Uniform	F1-F2	0.02442		
7	R-F	0.00608	R1-R2	0.04093	0.04093	Tidak Bias
		Non-Uniform	F1-F2	0.00800		
8	R-F	0.04265	R1-R2	0.05093	0.05093	Tidak Bias
		Non-Uniform	F1-F2	0.04959		
9	R-F	0.03527	R1-R2	0.05042	0.05042	Tidak Bias
		Uniform	F1-F2	0.01971		
10	R-F	0.02456	R1-R2	0.02033	0.02480	Tidak Bias
		Non-Uniform	F1-F2	0.02480		
11	R-F	0.02649	R1-R2	0.01767	0.01767	Bias
		Non-Uniform	F1-F2	0.00597		
12	R-F	0.04050	R1-R2	0.05266	0.05266	Tidak Bias
		Non-Uniform	F1-F2	0.03990		
13	R-F	0.06070	R1-R2	0.08026	0.08026	Tidak Bias
		Non-Uniform	F1-F2	0.05495		
14	R-F	0.04208	R1-R2	0.07777	0.07777	Tidak Bias
		Non-Uniform	F1-F2	0.04251		
15	R-F	0.06513	R1-R2	0.04370	0.04863	Bias
		Non-Uniform	F1-F2	0.04863		
16	R-F	0.04653	R1-R2	0.01484	0.01984	Bias
		Non-Uniform	F1-F2	0.01984		
17	R-F	0.00922	R1-R2	0.02171	0.03820	Tidak Bias
		Non-Uniform	F1-F2	0.03820		
18	R-F	0.02622	R1-R2	0.04887	0.04887	Tidak Bias
		Non-Uniform	F1-F2	0.03543		
19	R-F	0.01531	R1-R2	0.02458	0.04286	Tidak Bias
		Non-Uniform	F1-F2	0.04286		
20	R-F	0.02936	R1-R2	0.02383	0.03221	Tidak Bias
		Non-Uniform	F1-F2	0.03221		
21	R-F	0.04159	R1-R2	0.04723	0.04723	Tidak Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
22	R-F	Non-Uniform	F1-F2	0.04233	0.07645	Tidak Bias
		0.03151	R1-R2	0.07645		
23	R-F	Non-Uniform	F1-F2	0.00973	0.04645	Tidak Bias
		0.03843	R1-R2	0.04645		
24	R-F	Non-Uniform	F1-F2	0.02855	0.02721	Tidak Bias
		0.02621	R1-R2	0.01129		
25	R-F	Non-Uniform	F1-F2	0.02721	0.03439	Tidak Bias
		0.00945	R1-R2	0.03439		
26	R-F	Non-Uniform	F1-F2	0.01152	0.03771	Tidak Bias
		0.03507	R1-R2	0.03771		
27	R-F	Non-Uniform	F1-F2	0.01894	0.04230	Tidak Bias
		0.03834	R1-R2	0.04230		
28	R-F	Non-Uniform	F1-F2	0.02913	0.02154	Bias
		0.05415	R1-R2	0.02154		
29	R-F	Non-Uniform	F1-F2	0.01443	0.05710	Tidak Bias
		0.02067	R1-R2	0.03479		
JUMLAH BIAS BUTIR					5	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	4.63394315	7.81472	Tidak Bias
2	5.468331165	7.81472	Tidak Bias
3	1.067898188	7.81472	Tidak Bias
4	0.214405324	7.81472	Tidak Bias
5	4.674982404	7.81472	Tidak Bias
6	0.214290435	7.81472	Tidak Bias
7	0.231650979	7.81472	Tidak Bias
8	4.619735565	7.81472	Tidak Bias
9	0.144493692	7.81472	Tidak Bias
10	5.701349731	7.81472	Tidak Bias
11	2.709909212	7.81472	Tidak Bias
12	4.000084049	7.81472	Tidak Bias
13	8.704381461	7.81472	Bias
14	3.988979804	7.81472	Tidak Bias
15	12.62909701	7.81472	Bias
16	6.464483003	7.81472	Tidak Bias
17	1.858786048	7.81472	Tidak Bias
18	2.713201285	7.81472	Tidak Bias
19	0.922644412	7.81472	Tidak Bias
20	2.215822955	7.81472	Tidak Bias
21	3.835595355	7.81472	Tidak Bias
22	2.80452103	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	5.766147237	7.81472	Tidak Bias
24	1.432416548	7.81472	Tidak Bias
25	0.46778142	7.81472	Tidak Bias
26	3.781082346	7.81472	Tidak Bias
27	5.249692086	7.81472	Tidak Bias
28	4.189604735	7.81472	Tidak Bias
29	1.205317791	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			2

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.20781	0.12796	-1.62405	Tidak Bias
2	0.21724	0.24674	0.88043	Tidak Bias
3	-0.09705	0.10881	-0.89191	Tidak Bias
4	0.07039	0.16657	0.42260	Tidak Bias
5	-0.10580	0.16638	-0.63588	Tidak Bias
6	-0.01218	0.07902	-0.15419	Tidak Bias
7	-0.01643	0.10700	-0.15357	Tidak Bias
8	0.30197	0.14383	2.09951	Bias
9	0.35178	0.96993	0.36269	Tidak Bias
10	0.24993	0.32983	0.75775	Tidak Bias
11	0.10867	0.17287	0.62861	Tidak Bias
12	0.03988	0.08055	0.49510	Tidak Bias
13	-0.11438	0.06921	-1.65268	Tidak Bias
14	-0.19086	0.09767	-1.95411	Tidak Bias
15	-0.19506	0.07349	-2.65429	Bias
16	-0.30707	0.24730	-1.24170	Tidak Bias
17	-0.00678	0.05376	-0.12616	Tidak Bias
18	-0.05267	0.17243	-0.30547	Tidak Bias
19	-0.02161	0.06180	-0.34968	Tidak Bias
20	-0.01582	0.12785	-0.12374	Tidak Bias
21	0.20673	0.11100	1.86238	Tidak Bias
22	-0.15508	0.17034	-0.91039	Tidak Bias
23	-0.12423	0.08812	-1.40984	Tidak Bias
24	-0.07654	0.08046	-0.95121	Tidak Bias
25	-0.03047	0.06781	-0.44943	Tidak Bias
26	0.18826	0.10042	1.87467	Tidak Bias
27	-0.10483	0.08100	-1.29421	Tidak Bias
28	0.28415	0.25674	1.10676	Tidak Bias
29	-0.18431	0.19162	-0.96183	Tidak Bias
JUMLAH BIAS BUTIR				2

REPLIKASI KE-23.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok		Indeks Subkelompok		Maksimum	Kesimpulan
1	R-F	0.02684	R1-R2	0.01950	0.02856	Tidak Bias
		Non-Uniform	F1-F2	0.02856		
2	R-F	0.03889	R1-R2	0.08113	0.08113	Tidak Bias
		Non-Uniform	F1-F2	0.03214		
3	R-F	0.03760	R1-R2	0.04210	0.05661	Tidak Bias
		Non-Uniform	F1-F2	0.05661		
4	R-F	0.02408	R1-R2	0.03147	0.03147	Tidak Bias
		Non-Uniform	F1-F2	0.02221		
5	R-F	0.05020	R1-R2	0.02727	0.03673	Bias
		Non-Uniform	F1-F2	0.03673		
6	R-F	0.03488	R1-R2	0.02603	0.04521	Tidak Bias
		Non-Uniform	F1-F2	0.04521		
7	R-F	0.06215	R1-R2	0.02259	0.03443	Bias
		Non-Uniform	F1-F2	0.03443		
8	R-F	0.01946	R1-R2	0.04510	0.05249	Tidak Bias
		Non-Uniform	F1-F2	0.05249		
9	R-F	0.03833	R1-R2	0.01260	0.04422	Tidak Bias
		Non-Uniform	F1-F2	0.04422		
10	R-F	0.00953	R1-R2	0.03157	0.03157	Tidak Bias
		Uniform	F1-F2	0.01008		
11	R-F	0.01624	R1-R2	0.03274	0.03274	Tidak Bias
		Non-Uniform	F1-F2	0.01466		
12	R-F	0.03543	R1-R2	0.00784	0.03525	Bias
		Non-Uniform	F1-F2	0.03525		
13	R-F	0.04916	R1-R2	0.06715	0.06715	Tidak Bias
		Non-Uniform	F1-F2	0.05178		
14	R-F	0.04207	R1-R2	0.02297	0.05279	Tidak Bias
		Non-Uniform	F1-F2	0.05279		
15	R-F	0.04230	R1-R2	0.03484	0.03484	Bias
		Non-Uniform	F1-F2	0.02165		
16	R-F	0.03950	R1-R2	0.02433	0.02433	Bias
		Non-Uniform	F1-F2	0.02274		
17	R-F	0.03353	R1-R2	0.04122	0.04749	Tidak Bias
		Non-Uniform	F1-F2	0.04749		
18	R-F	0.03175	R1-R2	0.03368	0.03368	Tidak Bias
		Non-Uniform	F1-F2	0.02495		
19	R-F	0.02915	R1-R2	0.04815	0.04815	Tidak Bias
		Non-Uniform	F1-F2	0.03669		
20	R-F	0.02983	R1-R2	0.06962	0.06962	Tidak Bias
		Non-Uniform	F1-F2	0.01886		
21	R-F	0.05657	R1-R2	0.04024	0.04024	Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Non-Uniform	F1-F2	0.03884	Tidak Bias
		0.00940	R1-R2	0.02691	
23	R-F	Non-Uniform	F1-F2	0.02142	Tidak Bias
		0.03513	R1-R2	0.03911	
24	R-F	Non-Uniform	F1-F2	0.03288	Bias
		0.04057	R1-R2	0.02239	
25	R-F	Non-Uniform	F1-F2	0.03789	Bias
		0.05407	R1-R2	0.05212	
26	R-F	Non-Uniform	F1-F2	0.01134	Tidak Bias
		0.03662	R1-R2	0.04497	
27	R-F	Non-Uniform	F1-F2	0.01490	Tidak Bias
		0.01065	R1-R2	0.03434	
28	R-F	Non-Uniform	F1-F2	0.04535	Bias
		0.04263	R1-R2	0.01102	
29	R-F	Non-Uniform	F1-F2	0.02039	Tidak Bias
		0.04486	R1-R2	0.03162	
JUMLAH BIAS BUTIR				9	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	2.684978299	7.81472	Tidak Bias
2	7.465605138	7.81472	Tidak Bias
3	4.339774023	7.81472	Tidak Bias
4	4.384951113	7.81472	Tidak Bias
5	8.209879556	7.81472	Bias
6	4.572775483	7.81472	Tidak Bias
7	17.40081006	7.81472	Bias
8	0.818716022	7.81472	Tidak Bias
9	5.901947319	7.81472	Tidak Bias
10	0.607396106	7.81472	Tidak Bias
11	-0.410943489	7.81472	Tidak Bias
12	4.892982352	7.81472	Tidak Bias
13	3.222782044	7.81472	Tidak Bias
14	6.562692918	7.81472	Tidak Bias
15	5.186350866	7.81472	Tidak Bias
16	6.189837575	7.81472	Tidak Bias
17	2.758270000	7.81472	Tidak Bias
18	4.83655946	7.81472	Tidak Bias
19	4.167281186	7.81472	Tidak Bias
20	3.159681893	7.81472	Tidak Bias
21	9.541469068	7.81472	Bias
22	0.27058346	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	2.758011165	7.81472	Tidak Bias
24	6.973580142	7.81472	Tidak Bias
25	11.57498606	7.81472	Bias
26	3.267755524	7.81472	Tidak Bias
27	0.453848036	7.81472	Tidak Bias
28	4.352897418	7.81472	Tidak Bias
29	7.775475069	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			4

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.16324	0.16527	-0.98769	Tidak Bias
2	-0.23091	0.23200	-0.99529	Tidak Bias
3	-0.11268	0.10393	-1.08423	Tidak Bias
4	-0.00393	0.17077	-0.02299	Tidak Bias
5	-0.04472	0.18401	-0.24301	Tidak Bias
6	0.15104	0.07973	1.89451	Tidak Bias
7	0.15607	0.10500	1.48633	Tidak Bias
8	0.10602	0.14342	0.73925	Tidak Bias
9	-0.24837	1.01646	-0.24435	Tidak Bias
10	0.03337	0.34854	0.09573	Tidak Bias
11	-0.10569	0.14605	-0.72365	Tidak Bias
12	0.13106	0.08425	1.55551	Tidak Bias
13	0.09766	0.07783	1.25468	Tidak Bias
14	-0.12785	0.09818	-1.30220	Tidak Bias
15	-0.01805	0.07923	-0.22785	Tidak Bias
16	-0.31390	0.29515	-1.06352	Tidak Bias
17	0.07195	0.06068	1.18573	Tidak Bias
18	0.07368	0.15887	0.46378	Tidak Bias
19	0.04398	0.06751	0.65144	Tidak Bias
20	0.11896	0.13493	0.88164	Tidak Bias
21	0.24397	0.10976	2.22278	Bias
22	0.05439	0.17064	0.31877	Tidak Bias
23	-0.03049	0.09862	-0.30920	Tidak Bias
24	-0.04017	0.08495	-0.47290	Tidak Bias
25	0.10433	0.07695	1.35571	Tidak Bias
26	-0.02586	0.09631	-0.26854	Tidak Bias
27	0.00624	0.08105	0.07704	Tidak Bias
28	0.19807	0.27461	0.72129	Tidak Bias
29	-0.12492	0.18898	-0.66105	Tidak Bias
JUMLAH BIAS BUTIR				1

REPLIKASI KE-24.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.04381	R1-R2	0.04965	0.06285	Tidak Bias
		Non-Uniform	F1-F2	0.06285		
2	R-F	0.03670	R1-R2	0.01822	0.01822	Bias
		Uniform	F1-F2	0.01234		
3	R-F	0.04547	R1-R2	0.02145	0.04060	Bias
		Non-Uniform	F1-F2	0.04060		
4	R-F	0.04762	R1-R2	0.03190	0.05751	Tidak Bias
		Uniform	F1-F2	0.05751		
5	R-F	0.03057	R1-R2	0.05300	0.05300	Tidak Bias
		Non-Uniform	F1-F2	0.01746		
6	R-F	0.03288	R1-R2	0.04102	0.04102	Tidak Bias
		Uniform	F1-F2	0.03450		
7	R-F	0.07848	R1-R2	0.03182	0.03773	Bias
		Uniform	F1-F2	0.03773		
8	R-F	0.03272	R1-R2	0.02263	0.02263	Bias
		Non-Uniform	F1-F2	0.01622		
9	R-F	0.03969	R1-R2	0.04138	0.04138	Tidak Bias
		Uniform	F1-F2	0.03152		
10	R-F	0.02396	R1-R2	0.02057	0.02057	Bias
		Non-Uniform	F1-F2	0.00311		
11	R-F	0.03536	R1-R2	0.03107	0.04101	Tidak Bias
		Non-Uniform	F1-F2	0.04101		
12	R-F	0.04322	R1-R2	0.06631	0.06631	Tidak Bias
		Uniform	F1-F2	0.02637		
13	R-F	0.05431	R1-R2	0.01681	0.02654	Bias
		Non-Uniform	F1-F2	0.02654		
14	R-F	0.02235	R1-R2	0.04214	0.04214	Tidak Bias
		Uniform	F1-F2	0.02758		
15	R-F	0.01625	R1-R2	0.02036	0.10814	Tidak Bias
		Non-Uniform	F1-F2	0.10814		
16	R-F	0.06442	R1-R2	0.01185	0.01185	Bias
		Non-Uniform	F1-F2	0.00703		
17	R-F	0.03814	R1-R2	0.04926	0.07799	Tidak Bias
		Non-Uniform	F1-F2	0.07799		
18	R-F	0.02511	R1-R2	0.02533	0.02533	Tidak Bias
		Uniform	F1-F2	0.01640		
19	R-F	0.02185	R1-R2	0.04521	0.04521	Tidak Bias
		Non-Uniform	F1-F2	0.03990		
20	R-F	0.04188	R1-R2	0.03002	0.03604	Bias
		Non-Uniform	F1-F2	0.03604		
21	R-F	0.02503	R1-R2	0.01181	0.02106	Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Uniform	F1-F2	0.02106	Tidak Bias
		0.02320	R1-R2	0.02519	
23	R-F	Non-Uniform	F1-F2	0.04792	Tidak Bias
		0.02305	R1-R2	0.05824	
24	R-F	Non-Uniform	F1-F2	0.04732	Tidak Bias
		0.02016	R1-R2	0.02067	
25	R-F	Uniform	F1-F2	0.07047	Tidak Bias
		0.02071	R1-R2	0.07498	
26	R-F	Uniform	F1-F2	0.06680	Tidak Bias
		0.01448	R1-R2	0.01503	
27	R-F	Non-Uniform	F1-F2	0.00945	Tidak Bias
		0.01518	R1-R2	0.04808	
28	R-F	Non-Uniform	F1-F2	0.03270	Bias
		0.05591	R1-R2	0.00521	
29	R-F	Non-Uniform	F1-F2	0.04746	Bias
		0.02600	R1-R2	0.02427	
JUMLAH BIAS BUTIR				11	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	6.402139671	7.81472	Tidak Bias
2	6.709106792	7.81472	Tidak Bias
3	8.106564771	7.81472	Bias
4	11.40175165	7.81472	Bias
5	2.53575847	7.81472	Tidak Bias
6	2.565808845	7.81472	Tidak Bias
7	25.53370516	7.81472	Bias
8	3.266621991	7.81472	Tidak Bias
9	2.704543518	7.81472	Tidak Bias
10	2.72988461	7.81472	Tidak Bias
11	6.876448708	7.81472	Tidak Bias
12	8.013849104	7.81472	Bias
13	6.035885667	7.81472	Tidak Bias
14	1.086161663	7.81472	Tidak Bias
15	0.503083233	7.81472	Tidak Bias
16	10.60918439	7.81472	Bias
17	28.1990624	7.81472	Bias
18	3.279340679	7.81472	Tidak Bias
19	1.694089228	7.81472	Tidak Bias
20	5.60972395	7.81472	Tidak Bias
21	3.824280687	7.81472	Tidak Bias
22	1.697829345	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	0.847054984	7.81472	Tidak Bias
24	2.675952992	7.81472	Tidak Bias
25	1.279914331	7.81472	Tidak Bias
26	0.734985974	7.81472	Tidak Bias
27	0.474826841	7.81472	Tidak Bias
28	6.95775101	7.81472	Tidak Bias
29	2.95174786	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			6

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.23723	0.13770	-1.72282	Tidak Bias
2	0.24272	0.25875	0.93805	Tidak Bias
3	-0.15563	0.11656	-1.33526	Tidak Bias
4	0.18554	0.16667	1.11321	Tidak Bias
5	-0.01368	0.16887	-0.08102	Tidak Bias
6	0.08717	0.08275	1.05336	Tidak Bias
7	0.26215	0.11381	2.30345	Bias
8	0.00841	0.19248	0.04371	Tidak Bias
9	-0.07327	1.74758	-0.04193	Tidak Bias
10	-0.42522	0.36806	-1.15528	Tidak Bias
11	0.20253	0.18385	1.10162	Tidak Bias
12	0.13871	0.08677	1.59849	Tidak Bias
13	-0.04583	0.07254	-0.63180	Tidak Bias
14	0.01353	0.10788	0.12540	Tidak Bias
15	-0.04457	0.08049	-0.55373	Tidak Bias
16	-0.42921	0.25602	-1.67650	Tidak Bias
17	0.02048	0.05642	0.36299	Tidak Bias
18	0.12246	0.16932	0.72328	Tidak Bias
19	0.02571	0.06911	0.37194	Tidak Bias
20	0.03158	0.14671	0.21523	Tidak Bias
21	0.06859	0.11333	0.60526	Tidak Bias
22	-0.15639	0.15381	-1.01677	Tidak Bias
23	-0.05952	0.10112	-0.58865	Tidak Bias
24	0.03465	0.07680	0.45113	Tidak Bias
25	0.04301	0.07990	0.53832	Tidak Bias
26	0.07187	0.10779	0.66676	Tidak Bias
27	0.00205	0.07650	0.02678	Tidak Bias
28	0.30284	0.29410	1.02971	Tidak Bias
29	-0.22343	0.17382	-1.28544	Tidak Bias
JUMLAH BIAS BUTIR				1

REPLIKASI KE-25.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.03439	R1-R2	0.02899	0.05183	Tidak Bias
		Non-Uniform	F1-F2	0.05183		
2	R-F	0.02580	R1-R2	0.02431	0.02431	Bias
		Uniform	F1-F2	0.02276		
3	R-F	0.03922	R1-R2	0.03102	0.03102	Bias
		Non-Uniform	F1-F2	0.00884		
4	R-F	0.02027	R1-R2	0.04985	0.04985	Tidak Bias
		Non-Uniform	F1-F2	0.00902		
5	R-F	0.03442	R1-R2	0.02449	0.06149	Tidak Bias
		Non-Uniform	F1-F2	0.06149		
6	R-F	0.02551	R1-R2	0.07159	0.07159	Tidak Bias
		Non-Uniform	F1-F2	0.00951		
7	R-F	0.03792	R1-R2	0.01890	0.04853	Tidak Bias
		Uniform	F1-F2	0.04853		
8	R-F	0.03534	R1-R2	0.07951	0.07951	Tidak Bias
		Non-Uniform	F1-F2	0.04958		
9	R-F	0.03622	R1-R2	0.03576	0.03576	Bias
		Uniform	F1-F2	0.01820		
10	R-F	0.02095	R1-R2	0.04313	0.04313	Tidak Bias
		Non-Uniform	F1-F2	0.01228		
11	R-F	0.01630	R1-R2	0.01835	0.04950	Tidak Bias
		Non-Uniform	F1-F2	0.04950		
12	R-F	0.02200	R1-R2	0.03223	0.03223	Tidak Bias
		Non-Uniform	F1-F2	0.01203		
13	R-F	0.06177	R1-R2	0.07424	0.07424	Tidak Bias
		Non-Uniform	F1-F2	0.01673		
14	R-F	0.02552	R1-R2	0.03283	0.06278	Tidak Bias
		Non-Uniform	F1-F2	0.06278		
15	R-F	0.03221	R1-R2	0.04643	0.04643	Tidak Bias
		Non-Uniform	F1-F2	0.01293		
16	R-F	0.04054	R1-R2	0.06262	0.06262	Tidak Bias
		Non-Uniform	F1-F2	0.04792		
17	R-F	0.06819	R1-R2	0.08323	0.08323	Tidak Bias
		Non-Uniform	F1-F2	0.04065		
18	R-F	0.04422	R1-R2	0.00684	0.02640	Bias
		Non-Uniform	F1-F2	0.02640		
19	R-F	0.01431	R1-R2	0.03146	0.03146	Tidak Bias
		Non-Uniform	F1-F2	0.01112		
20	R-F	0.03327	R1-R2	0.02965	0.04628	Tidak Bias
		Non-Uniform	F1-F2	0.04628		
21	R-F	0.03796	R1-R2	0.05595	0.05595	Tidak Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Uniform	F1-F2	0.03128	Tidak Bias
		0.01886	R1-R2	0.03590	
23	R-F	Non-Uniform	F1-F2	0.03131	Tidak Bias
		0.02291	R1-R2	0.02232	
24	R-F	Non-Uniform	F1-F2	0.04966	Bias
		0.05145	R1-R2	0.02667	
25	R-F	Non-Uniform	F1-F2	0.01097	Tidak Bias
		0.01442	R1-R2	0.02772	
26	R-F	Non-Uniform	F1-F2	0.03808	Tidak Bias
		0.02260	R1-R2	0.06049	
27	R-F	Non-Uniform	F1-F2	0.03286	Tidak Bias
		0.01645	R1-R2	0.01292	
28	R-F	Non-Uniform	F1-F2	0.03188	Tidak Bias
		0.03150	R1-R2	0.04130	
29	R-F	Uniform	F1-F2	0.02975	Bias
		0.03463	R1-R2	0.00693	
JUMLAH BIAS BUTIR				6	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	2.771147661	7.81472	Tidak Bias
2	2.739525933	7.81472	Tidak Bias
3	3.039903592	7.81472	Tidak Bias
4	1.522873006	7.81472	Tidak Bias
5	3.947328772	7.81472	Tidak Bias
6	1.238547811	7.81472	Tidak Bias
7	6.426456497	7.81472	Tidak Bias
8	4.735434263	7.81472	Tidak Bias
9	2.428484786	7.81472	Tidak Bias
10	5.406749674	7.81472	Tidak Bias
11	-1726.240364	7.81472	Tidak Bias
12	1.295316823	7.81472	Tidak Bias
13	12.1136515	7.81472	Bias
14	2.471197835	7.81472	Tidak Bias
15	2.80852558	7.81472	Tidak Bias
16	4.141603707	7.81472	Tidak Bias
17	11.25479358	7.81472	Bias
18	4.041535124	7.81472	Tidak Bias
19	0.545731936	7.81472	Tidak Bias
20	4.013386883	7.81472	Tidak Bias
21	5.28159558	7.81472	Tidak Bias
22	1.283952106	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	1.408700136	7.81472	Tidak Bias
24	6.254405741	7.81472	Tidak Bias
25	0.439438296	7.81472	Tidak Bias
26	1.43394377	7.81472	Tidak Bias
27	0.814929778	7.81472	Tidak Bias
28	3.124225323	7.81472	Tidak Bias
29	3.730509896	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			2

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.24355	0.14777	-1.64815	Tidak Bias
2	0.21663	0.25371	0.85385	Tidak Bias
3	-0.16056	0.11221	-1.43086	Tidak Bias
4	0.12101	0.17335	0.69806	Tidak Bias
5	0.09216	0.17138	0.53774	Tidak Bias
6	0.07438	0.08448	0.88046	Tidak Bias
7	0.13687	0.10676	1.28202	Tidak Bias
8	0.17127	0.17255	0.99260	Tidak Bias
9	-0.10784	1.46344	-0.07369	Tidak Bias
10	-0.29937	0.33397	-0.89640	Tidak Bias
11	-0.07079	0.15989	-0.44276	Tidak Bias
12	0.08121	0.08483	0.95727	Tidak Bias
13	-0.05054	0.07454	-0.67794	Tidak Bias
14	-0.11446	0.09862	-1.16055	Tidak Bias
15	-0.08495	0.07948	-1.06886	Tidak Bias
16	-0.23658	0.25852	-0.91511	Tidak Bias
17	0.14958	0.05630	2.65703	Bias
18	0.23197	0.18923	1.22587	Tidak Bias
19	-0.05022	0.07273	-0.69054	Tidak Bias
20	0.10215	0.12506	0.81687	Tidak Bias
21	0.04720	0.12636	0.37352	Tidak Bias
22	-0.07556	0.15331	-0.49287	Tidak Bias
23	-0.10805	0.10462	-1.03279	Tidak Bias
24	-0.03876	0.08341	-0.46477	Tidak Bias
25	-0.02056	0.07967	-0.25806	Tidak Bias
26	0.12005	0.10389	1.15554	Tidak Bias
27	-0.02558	0.07859	-0.32546	Tidak Bias
28	0.11138	0.24040	0.46331	Tidak Bias
29	0.03151	0.17604	0.17898	Tidak Bias
JUMLAH BIAS BUTIR				1

REPLIKASI KE-26.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok		Indeks Subkelompok		Maksimum	Kesimpulan
1	R-F	0.05184	R1-R2	0.06538	0.06538	Tidak Bias
		Non-Uniform	F1-F2	0.00940		
2	R-F	0.03276	R1-R2	0.01174	0.03300	Tidak Bias
		Uniform	F1-F2	0.03300		
3	R-F	0.04731	R1-R2	0.04311	0.05437	Tidak Bias
		Non-Uniform	F1-F2	0.05437		
4	R-F	0.02648	R1-R2	0.01559	0.02081	Bias
		Non-Uniform	F1-F2	0.02081		
5	R-F	0.02960	R1-R2	0.04761	0.04859	Tidak Bias
		Non-Uniform	F1-F2	0.04859		
6	R-F	0.05340	R1-R2	0.02875	0.03475	Bias
		Non-Uniform	F1-F2	0.03475		
7	R-F	0.04576	R1-R2	0.03078	0.03078	Bias
		Non-Uniform	F1-F2	0.02792		
8	R-F	0.02111	R1-R2	0.04613	0.04613	Tidak Bias
		Uniform	F1-F2	0.03487		
9	R-F	0.05488	R1-R2	0.02774	0.05399	Bias
		Uniform	F1-F2	0.05399		
10	R-F	0.00740	R1-R2	0.01729	0.07518	Tidak Bias
		Uniform	F1-F2	0.07518		
11	R-F	0.05083	R1-R2	0.01492	0.01545	Bias
		Non-Uniform	F1-F2	0.01545		
12	R-F	0.02599	R1-R2	0.03590	0.08271	Tidak Bias
		Uniform	F1-F2	0.08271		
13	R-F	0.07829	R1-R2	0.05173	0.05173	Bias
		Non-Uniform	F1-F2	0.03119		
14	R-F	0.04335	R1-R2	0.04177	0.05360	Tidak Bias
		Non-Uniform	F1-F2	0.05360		
15	R-F	0.01642	R1-R2	0.05295	0.05295	Tidak Bias
		Non-Uniform	F1-F2	0.02209		
16	R-F	0.02564	R1-R2	0.00941	0.01869	Bias
		Non-Uniform	F1-F2	0.01869		
17	R-F	0.03315	R1-R2	0.02947	0.02947	Bias
		Non-Uniform	F1-F2	0.02617		
18	R-F	0.00747	R1-R2	0.01150	0.04385	Tidak Bias
		Uniform	F1-F2	0.04385		
19	R-F	0.04398	R1-R2	0.01565	0.02062	Bias
		Non-Uniform	F1-F2	0.02062		
20	R-F	0.01700	R1-R2	0.04884	0.04884	Tidak Bias
		Non-Uniform	F1-F2	0.03967		
21	R-F	0.03394	R1-R2	0.01905	0.04850	Tidak Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Non-Uniform	F1-F2	0.04850	Tidak Bias
		0.01176	R1-R2	0.03386	
23	R-F	Uniform	F1-F2	0.01364	Tidak Bias
		0.04306	R1-R2	0.06506	
24	R-F	Non-Uniform	F1-F2	0.04741	Bias
		0.03257	R1-R2	0.01381	
25	R-F	Non-Uniform	F1-F2	0.01079	Tidak Bias
		0.01691	R1-R2	0.04399	
26	R-F	Non-Uniform	F1-F2	0.02475	Bias
		0.03089	R1-R2	0.03086	
27	R-F	Non-Uniform	F1-F2	0.03058	Bias
		0.06104	R1-R2	0.01196	
28	R-F	Non-Uniform	F1-F2	0.03018	Bias
		0.06323	R1-R2	0.03470	
29	R-F	Non-Uniform	F1-F2	0.00888	Bias
		0.02544	R1-R2	0.02171	
JUMLAH BIAS BUTIR				14	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	9.455212763	7.81472	Bias
2	5.308856354	7.81472	Tidak Bias
3	4.924304573	7.81472	Tidak Bias
4	1.879668045	7.81472	Tidak Bias
5	2.011032635	7.81472	Tidak Bias
6	4.154701088	7.81472	Tidak Bias
7	7.483341201	7.81472	Tidak Bias
8	1.581805534	7.81472	Tidak Bias
9	6.00117541	7.81472	Tidak Bias
10	0.436805037	7.81472	Tidak Bias
11	17.83111168	7.81472	Bias
12	3.075025186	7.81472	Tidak Bias
13	16.65477337	7.81472	Bias
14	6.109275574	7.81472	Tidak Bias
15	0.893143154	7.81472	Tidak Bias
16	2.44366627	7.81472	Tidak Bias
17	3.098543756	7.81472	Tidak Bias
18	0.778857673	7.81472	Tidak Bias
19	8.876573078	7.81472	Bias
20	0.645796576	7.81472	Tidak Bias
21	3.099920298	7.81472	Tidak Bias
22	0.794809715	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	7.667313744	7.81472	Tidak Bias
24	4.784288133	7.81472	Tidak Bias
25	0.84736395	7.81472	Tidak Bias
26	3.488088727	7.81472	Tidak Bias
27	9.818726767	7.81472	Bias
28	8.301806494	7.81472	Bias
29	1.293551203	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			6

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.29458	0.13049	-2.25760	Bias
2	0.17381	0.25576	0.67961	Tidak Bias
3	-0.21130	0.11151	-1.89499	Tidak Bias
4	0.12609	0.17494	0.72077	Tidak Bias
5	-0.02123	0.20881	-0.10169	Tidak Bias
6	0.08750	0.08134	1.07576	Tidak Bias
7	0.12057	0.11790	1.02266	Tidak Bias
8	-0.04247	0.17029	-0.24943	Tidak Bias
9	0.28741	4.65974	0.06168	Tidak Bias
10	-0.13021	0.32211	-0.40423	Tidak Bias
11	0.35994	0.17781	2.02431	Bias
12	0.04473	0.07641	0.58545	Tidak Bias
13	-0.20253	0.07214	-2.80748	Bias
14	-0.23048	0.10801	-2.13398	Bias
15	-0.06032	0.06993	-0.86270	Tidak Bias
16	-0.16610	0.24076	-0.68991	Tidak Bias
17	-0.07553	0.05646	-1.33761	Tidak Bias
18	0.01235	0.17991	0.06866	Tidak Bias
19	-0.12711	0.06852	-1.85491	Tidak Bias
20	-0.09397	0.13847	-0.67862	Tidak Bias
21	-0.07967	0.11873	-0.67107	Tidak Bias
22	0.02261	0.15634	0.14462	Tidak Bias
23	-0.06456	0.08483	-0.76103	Tidak Bias
24	-0.14419	0.08389	-1.71895	Tidak Bias
25	-0.00471	0.06935	-0.06785	Tidak Bias
26	0.12216	0.11065	1.10406	Tidak Bias
27	-0.06471	0.08017	-0.80717	Tidak Bias
28	0.51478	0.29127	1.76734	Tidak Bias
29	0.14171	0.17612	0.80464	Tidak Bias
JUMLAH BIAS BUTIR				4

REPLIKASI KE-27.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok		Indeks Subkelompok		Maksimum	Kesimpulan
1	R-F	0.06916	R1-R2	0.02738	0.04047	Bias
		Non-Uniform	F1-F2	0.04047		
2	R-F	0.02231	R1-R2	0.02494	0.03389	Tidak Bias
		Non-Uniform	F1-F2	0.03389		
3	R-F	0.03768	R1-R2	0.04965	0.04965	Tidak Bias
		Non-Uniform	F1-F2	0.02381		
4	R-F	0.03959	R1-R2	0.03189	0.09444	Tidak Bias
		Non-Uniform	F1-F2	0.09444		
5	R-F	0.03736	R1-R2	0.03944	0.08526	Tidak Bias
		Non-Uniform	F1-F2	0.08526		
6	R-F	0.03296	R1-R2	0.06691	0.06691	Tidak Bias
		Non-Uniform	F1-F2	0.02991		
7	R-F	0.02884	R1-R2	0.09074	0.09074	Tidak Bias
		Non-Uniform	F1-F2	0.03752		
8	R-F	0.03110	R1-R2	0.06636	0.06636	Tidak Bias
		Non-Uniform	F1-F2	0.03761		
9	R-F	0.04425	R1-R2	0.07922	0.07922	Tidak Bias
		Non-Uniform	F1-F2	0.03771		
10	R-F	0.03123	R1-R2	0.00922	0.03272	Tidak Bias
		Non-Uniform	F1-F2	0.03272		
11	R-F	0.05821	R1-R2	0.03023	0.04405	Bias
		Non-Uniform	F1-F2	0.04405		
12	R-F	0.03890	R1-R2	0.04776	0.05348	Tidak Bias
		Non-Uniform	F1-F2	0.05348		
13	R-F	0.11098	R1-R2	0.08221	0.08221	Bias
		Non-Uniform	F1-F2	0.05584		
14	R-F	0.10012	R1-R2	0.03617	0.10642	Tidak Bias
		Non-Uniform	F1-F2	0.10642		
15	R-F	0.07898	R1-R2	0.05218	0.05533	Bias
		Non-Uniform	F1-F2	0.05533		
16	R-F	0.01570	R1-R2	0.04166	0.04166	Tidak Bias
		Non-Uniform	F1-F2	0.01922		
17	R-F	0.09129	R1-R2	0.10644	0.10644	Tidak Bias
		Non-Uniform	F1-F2	0.03668		
18	R-F	0.04759	R1-R2	0.04101	0.04101	Bias
		Non-Uniform	F1-F2	0.02884		
19	R-F	0.07405	R1-R2	0.02005	0.04015	Bias
		Non-Uniform	F1-F2	0.04015		
20	R-F	0.01935	R1-R2	0.04098	0.04098	Tidak Bias
		Non-Uniform	F1-F2	0.03113		
21	R-F	0.03899	R1-R2	0.06030	0.07481	Tidak Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Non-Uniform	F1-F2	0.07481	Bias
		0.05166	R1-R2	0.03674	
23	R-F	Non-Uniform	F1-F2	0.03401	Tidak Bias
		0.02490	R1-R2	0.03854	
24	R-F	Uniform	F1-F2	0.02791	Bias
		0.06407	R1-R2	0.03836	
25	R-F	Non-Uniform	F1-F2	0.03578	Tidak Bias
		0.05531	R1-R2	0.06086	
26	R-F	Non-Uniform	F1-F2	0.05633	Bias
		0.04543	R1-R2	0.04532	
27	R-F	Non-Uniform	F1-F2	0.04230	Tidak Bias
		0.03593	R1-R2	0.07882	
28	R-F	Non-Uniform	F1-F2	0.05484	Bias
		0.05491	R1-R2	0.02997	
29	R-F	Non-Uniform	F1-F2	0.02440	Tidak Bias
		0.03856	R1-R2	0.02817	
JUMLAH BIAS BUTIR				10	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	13.51379779	7.81472	Bias
2	1.421381241	7.81472	Tidak Bias
3	5.452383806	7.81472	Tidak Bias
4	3.905014404	7.81472	Tidak Bias
5	4.294273936	7.81472	Tidak Bias
6	3.372527814	7.81472	Tidak Bias
7	2.20893515	7.81472	Tidak Bias
8	2.387364222	7.81472	Tidak Bias
9	0.65764312	7.81472	Tidak Bias
10	7.443515348	7.81472	Tidak Bias
11	21.70485465	7.81472	Bias
12	5.249355528	7.81472	Tidak Bias
13	37.44457955	7.81472	Bias
14	26.98237178	7.81472	Bias
15	16.87135833	7.81472	Bias
16	0.958119591	7.81472	Tidak Bias
17	24.52847941	7.81472	Bias
18	5.801288996	7.81472	Tidak Bias
19	16.04566697	7.81472	Bias
20	1.155208724	7.81472	Tidak Bias
21	6.089037012	7.81472	Tidak Bias
22	9.877662491	7.81472	Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	5.716254098	7.81472	Tidak Bias
24	16.27901571	7.81472	Bias
25	6.131114105	7.81472	Tidak Bias
26	5.484580698	7.81472	Tidak Bias
27	4.121535565	7.81472	Tidak Bias
28	5.985837628	7.81472	Tidak Bias
29	5.560660247	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			9

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.40632	0.15167	-2.67907	Bias
2	0.22622	0.28332	0.79847	Tidak Bias
3	-0.15826	0.12016	-1.31712	Tidak Bias
4	0.21774	0.18781	1.15936	Tidak Bias
5	-0.10604	0.17487	-0.60638	Tidak Bias
6	-0.04240	0.08169	-0.51902	Tidak Bias
7	0.11244	0.11769	0.95536	Tidak Bias
8	0.21032	0.14361	1.46447	Tidak Bias
9	0.94140	5.43986	0.17306	Tidak Bias
10	0.67469	0.33229	2.03043	Bias
11	0.45382	0.17037	2.66364	Bias
12	-0.07946	0.08322	-0.95478	Tidak Bias
13	-0.26680	0.06777	-3.93701	Bias
14	-0.43109	0.10118	-4.26080	Bias
15	-0.18452	0.07414	-2.48895	Bias
16	-0.09362	0.26907	-0.34796	Tidak Bias
17	-0.22420	0.06094	-3.67929	Bias
18	0.21312	0.17157	1.24212	Tidak Bias
19	-0.15933	0.06458	-2.46720	Bias
20	-0.04646	0.14134	-0.32874	Tidak Bias
21	-0.05091	0.10107	-0.50370	Tidak Bias
22	-0.09277	0.15508	-0.59822	Tidak Bias
23	-0.05914	0.09594	-0.61643	Tidak Bias
24	-0.24689	0.08026	-3.07606	Bias
25	-0.12958	0.07381	-1.75562	Tidak Bias
26	-0.09211	0.10154	-0.90712	Tidak Bias
27	-0.05078	0.07412	-0.68512	Tidak Bias
28	0.25403	0.27842	0.91239	Tidak Bias
29	-0.38310	0.16859	-2.27245	Bias
JUMLAH BIAS BUTIR				10

REPLIKASI KE-28.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.05245	R1-R2	0.02703	0.06385	Tidak Bias
		Non-Uniform	F1-F2	0.06385		
2	R-F	0.05104	R1-R2	0.01779	0.03001	Bias
		Non-Uniform	F1-F2	0.03001		
3	R-F	0.03862	R1-R2	0.07218	0.07218	Tidak Bias
		Non-Uniform	F1-F2	0.01076		
4	R-F	0.01836	R1-R2	0.01209	0.03383	Tidak Bias
		Non-Uniform	F1-F2	0.03383		
5	R-F	0.04339	R1-R2	0.03834	0.03834	Bias
		Non-Uniform	F1-F2	0.01436		
6	R-F	0.03372	R1-R2	0.01582	0.05437	Tidak Bias
		Uniform	F1-F2	0.05437		
7	R-F	0.04247	R1-R2	0.00346	0.03491	Bias
		Non-Uniform	F1-F2	0.03491		
8	R-F	0.04588	R1-R2	0.02394	0.02394	Bias
		Non-Uniform	F1-F2	0.01418		
9	R-F	0.04189	R1-R2	0.03725	0.03725	Bias
		Uniform	F1-F2	0.02110		
10	R-F	0.01167	R1-R2	0.05724	0.05724	Tidak Bias
		Non-Uniform	F1-F2	0.03141		
11	R-F	0.01373	R1-R2	0.01981	0.04011	Tidak Bias
		Uniform	F1-F2	0.04011		
12	R-F	0.04639	R1-R2	0.04320	0.09191	Tidak Bias
		Uniform	F1-F2	0.09191		
13	R-F	0.04637	R1-R2	0.04575	0.04575	Bias
		Uniform	F1-F2	0.02133		
14	R-F	0.03178	R1-R2	0.01391	0.06863	Tidak Bias
		Uniform	F1-F2	0.06863		
15	R-F	0.02587	R1-R2	0.02366	0.07474	Tidak Bias
		Non-Uniform	F1-F2	0.07474		
16	R-F	0.05301	R1-R2	0.03799	0.03799	Bias
		Non-Uniform	F1-F2	0.01645		
17	R-F	0.03044	R1-R2	0.02631	0.06938	Tidak Bias
		Uniform	F1-F2	0.06938		
18	R-F	0.01870	R1-R2	0.02237	0.03736	Tidak Bias
		Uniform	F1-F2	0.03736		
19	R-F	0.02124	R1-R2	0.02769	0.03519	Tidak Bias
		Non-Uniform	F1-F2	0.03519		
20	R-F	0.04432	R1-R2	0.04308	0.04308	Bias
		Non-Uniform	F1-F2	0.03059		
21	R-F	0.04962	R1-R2	0.03505	0.03505	Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Uniform	F1-F2	0.01882	Bias
		0.05104	R1-R2	0.03039	
		Non-Uniform	F1-F2	0.02744	
23	R-F	0.02219	R1-R2	0.06734	Tidak Bias
		Non-Uniform	F1-F2	0.04558	
24	R-F	0.03565	R1-R2	0.05312	Tidak Bias
		Uniform	F1-F2	0.03201	
25	R-F	0.05178	R1-R2	0.04036	Bias
		Non-Uniform	F1-F2	0.03568	
26	R-F	0.04751	R1-R2	0.02176	Tidak Bias
		Non-Uniform	F1-F2	0.07344	
27	R-F	0.01564	R1-R2	0.02880	Tidak Bias
		Uniform	F1-F2	0.04151	
28	R-F	0.03894	R1-R2	0.01841	Bias
		Uniform	F1-F2	0.00354	
29	R-F	0.04392	R1-R2	0.02495	Bias
		Non-Uniform	F1-F2	0.01732	
JUMLAH BIAS BUTIR				13	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	6.778168971	7.81472	Tidak Bias
2	19.70991436	7.81472	Bias
3	7.171764643	7.81472	Tidak Bias
4	0.747169404	7.81472	Tidak Bias
5	7.446154478	7.81472	Tidak Bias
6	5.416582451	7.81472	Tidak Bias
7	8.970693459	7.81472	Bias
8	7.542436317	7.81472	Tidak Bias
9	0.30341605	7.81472	Tidak Bias
10	1.407627045	7.81472	Tidak Bias
11	3.805375426	7.81472	Tidak Bias
12	25.88063607	7.81472	Bias
13	10.3156246	7.81472	Bias
14	2.508066812	7.81472	Tidak Bias
15	2.817534016	7.81472	Tidak Bias
16	8.58199873	7.81472	Bias
17	8.716647988	7.81472	Bias
18	3.375126927	7.81472	Tidak Bias
19	0.822204397	7.81472	Tidak Bias
20	9.229472863	7.81472	Bias
21	10.29107846	7.81472	Bias
22	5.575741728	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	3.162127722	7.81472	Tidak Bias
24	8.395658941	7.81472	Bias
25	12.07109292	7.81472	Bias
26	7.487406037	7.81472	Tidak Bias
27	0.796101827	7.81472	Tidak Bias
28	2.983063149	7.81472	Tidak Bias
29	10.30663732	7.81472	Bias
JUMLAH BIAS BUTIR			11

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.30848	0.13574	-2.27256	Bias
2	-0.00437	0.25894	-0.01686	Tidak Bias
3	-0.10244	0.11011	-0.93037	Tidak Bias
4	-0.07423	0.19617	-0.37837	Tidak Bias
5	0.12825	0.16622	0.77155	Tidak Bias
6	0.12754	0.07960	1.60229	Tidak Bias
7	0.08734	0.11014	0.79298	Tidak Bias
8	-0.13935	0.19361	-0.71974	Tidak Bias
9	-0.19752	1.03585	-0.19068	Tidak Bias
10	-0.16772	0.30646	-0.54728	Tidak Bias
11	0.05934	0.17134	0.34634	Tidak Bias
12	0.12440	0.07187	1.73092	Tidak Bias
13	0.01393	0.06698	0.20801	Tidak Bias
14	0.01479	0.10289	0.14373	Tidak Bias
15	0.01127	0.07472	0.15079	Tidak Bias
16	-0.29934	0.23366	-1.28111	Tidak Bias
17	0.08160	0.05767	1.41501	Tidak Bias
18	0.01124	0.17589	0.06392	Tidak Bias
19	0.01733	0.07220	0.24006	Tidak Bias
20	0.04165	0.13616	0.30587	Tidak Bias
21	0.14522	0.10976	1.32302	Tidak Bias
22	-0.09448	0.14896	-0.63427	Tidak Bias
23	-0.00064	0.09900	-0.00650	Tidak Bias
24	0.04467	0.08507	0.52517	Tidak Bias
25	0.12227	0.07831	1.56139	Tidak Bias
26	0.17996	0.10524	1.71001	Tidak Bias
27	-0.01074	0.08581	-0.12512	Tidak Bias
28	0.25525	0.28979	0.88081	Tidak Bias
29	-0.06675	0.15993	-0.41735	Tidak Bias
JUMLAH BIAS BUTIR				1

REPLIKASI KE-29.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.06159	R1-R2	0.02019	0.04254	Bias
		Non-Uniform	F1-F2	0.04254		
2	R-F	0.04013	R1-R2	0.02400	0.02400	Bias
		Non-Uniform	F1-F2	0.01191		
3	R-F	0.03684	R1-R2	0.01777	0.04233	Tidak Bias
		Non-Uniform	F1-F2	0.04233		
4	R-F	0.02922	R1-R2	0.01983	0.01983	Bias
		Non-Uniform	F1-F2	0.01898		
5	R-F	0.03541	R1-R2	0.03533	0.03533	Bias
		Non-Uniform	F1-F2	0.02829		
6	R-F	0.03430	R1-R2	0.01830	0.07847	Tidak Bias
		Uniform	F1-F2	0.07847		
7	R-F	0.04437	R1-R2	0.03864	0.05347	Tidak Bias
		Uniform	F1-F2	0.05347		
8	R-F	0.02067	R1-R2	0.04264	0.08198	Tidak Bias
		Uniform	F1-F2	0.08198		
9	R-F	0.05580	R1-R2	0.04528	0.04528	Bias
		Uniform	F1-F2	0.02760		
10	R-F	0.02259	R1-R2	0.02778	0.06609	Tidak Bias
		Non-Uniform	F1-F2	0.06609		
11	R-F	0.01043	R1-R2	0.01936	0.01936	Tidak Bias
		Non-Uniform	F1-F2	0.01838		
12	R-F	0.02591	R1-R2	0.05300	0.05300	Tidak Bias
		Non-Uniform	F1-F2	0.03130		
13	R-F	0.05484	R1-R2	0.08774	0.08774	Tidak Bias
		Uniform	F1-F2	0.07921		
14	R-F	0.02130	R1-R2	0.06741	0.06926	Tidak Bias
		Non-Uniform	F1-F2	0.06926		
15	R-F	0.01383	R1-R2	0.02443	0.04118	Tidak Bias
		Non-Uniform	F1-F2	0.04118		
16	R-F	0.04199	R1-R2	0.01949	0.01949	Bias
		Non-Uniform	F1-F2	0.00918		
17	R-F	0.02864	R1-R2	0.06785	0.08040	Tidak Bias
		Uniform	F1-F2	0.08040		
18	R-F	0.03026	R1-R2	0.03571	0.03571	Tidak Bias
		Uniform	F1-F2	0.01385		
19	R-F	0.03287	R1-R2	0.05292	0.05292	Tidak Bias
		Non-Uniform	F1-F2	0.03155		
20	R-F	0.02554	R1-R2	0.04560	0.09118	Tidak Bias
		Non-Uniform	F1-F2	0.09118		
21	R-F	0.05206	R1-R2	0.05379	0.05379	Tidak Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Non-Uniform	F1-F2	0.03462	Tidak Bias
		0.01046	R1-R2	0.02845	
23	R-F	Uniform	F1-F2	0.02667	Tidak Bias
		0.02502	R1-R2	0.08119	
24	R-F	Non-Uniform	F1-F2	0.05356	Bias
		0.03832	R1-R2	0.02666	
25	R-F	Non-Uniform	F1-F2	0.02984	Tidak Bias
		0.03543	R1-R2	0.06958	
26	R-F	Non-Uniform	F1-F2	0.03409	Tidak Bias
		0.02995	R1-R2	0.02656	
27	R-F	Non-Uniform	F1-F2	0.06753	Tidak Bias
		0.03065	R1-R2	0.03339	
28	R-F	Non-Uniform	F1-F2	0.05283	Bias
		0.02420	R1-R2	0.02046	
29	R-F	Uniform	F1-F2	0.02040	Tidak Bias
		0.01785	R1-R2	0.02457	
JUMLAH BIAS BUTIR				8	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	10.48260334	7.81472	Bias
2	8.248447716	7.81472	Bias
3	3.405696139	7.81472	Tidak Bias
4	4.065647716	7.81472	Tidak Bias
5	4.015764085	7.81472	Tidak Bias
6	7.393867007	7.81472	Tidak Bias
7	8.91033157	7.81472	Bias
8	1.361467888	7.81472	Tidak Bias
9	6.187628998	7.81472	Tidak Bias
10	1.34944178	7.81472	Tidak Bias
11	0.396504179	7.81472	Tidak Bias
12	2.580103755	7.81472	Tidak Bias
13	11.08137207	7.81472	Bias
14	1.662882786	7.81472	Tidak Bias
15	0.402712995	7.81472	Tidak Bias
16	4.919189608	7.81472	Tidak Bias
17	6.620848713	7.81472	Tidak Bias
18	4.873011889	7.81472	Tidak Bias
19	4.421385333	7.81472	Tidak Bias
20	2.181174365	7.81472	Tidak Bias
21	6.243914069	7.81472	Tidak Bias
22	0.352648676	7.81472	Tidak Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	1.114336448	7.81472	Tidak Bias
24	9.632880851	7.81472	Bias
25	4.888627202	7.81472	Tidak Bias
26	2.791871439	7.81472	Tidak Bias
27	1.854156738	7.81472	Tidak Bias
28	1.499498199	7.81472	Tidak Bias
29	1.253203246	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			5

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.45949	0.14356	-3.20059	Bias
2	0.13766	0.25981	0.52986	Tidak Bias
3	-0.21862	0.12420	-1.76030	Tidak Bias
4	0.00237	0.18826	0.01261	Tidak Bias
5	0.06257	0.16865	0.37100	Tidak Bias
6	0.12641	0.08052	1.56989	Tidak Bias
7	0.14701	0.11599	1.26741	Tidak Bias
8	0.08623	0.13474	0.63995	Tidak Bias
9	-0.14037	2.23406	-0.06283	Tidak Bias
10	-0.28971	0.31507	-0.91950	Tidak Bias
11	0.08499	0.17834	0.47657	Tidak Bias
12	0.08761	0.09016	0.97175	Tidak Bias
13	-0.00266	0.07192	-0.03697	Tidak Bias
14	-0.09657	0.10015	-0.96426	Tidak Bias
15	-0.04216	0.07819	-0.53922	Tidak Bias
16	-0.27813	0.25414	-1.09439	Tidak Bias
17	0.04192	0.05650	0.74188	Tidak Bias
18	0.10265	0.17656	0.58139	Tidak Bias
19	0.10084	0.06904	1.46051	Tidak Bias
20	0.04099	0.13437	0.30506	Tidak Bias
21	0.22318	0.10898	2.04786	Bias
22	-0.04689	0.15176	-0.30900	Tidak Bias
23	-0.08101	0.09742	-0.83158	Tidak Bias
24	0.03486	0.08319	0.41912	Tidak Bias
25	0.08786	0.07772	1.13043	Tidak Bias
26	0.12571	0.11065	1.13611	Tidak Bias
27	0.10042	0.08345	1.20332	Tidak Bias
28	0.11506	0.28379	0.40544	Tidak Bias
29	-0.05273	0.19439	-0.27128	Tidak Bias
JUMLAH BIAS BUTIR				2

REPLIKASI KE-30.

1. Metode Distribusi Sampling Empiris (DSE)

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan		
1	R-F	0.05776	R1-R2	0.01094	0.01443	Bias
		Non-Uniform	F1-F2	0.01443		
2	R-F	0.02583	R1-R2	0.02750	0.02750	Tidak Bias
		Uniform	F1-F2	0.01993		
3	R-F	0.02974	R1-R2	0.02220	0.03773	Tidak Bias
		Non-Uniform	F1-F2	0.03773		
4	R-F	0.03869	R1-R2	0.03285	0.03738	Bias
		Non-Uniform	F1-F2	0.03738		
5	R-F	0.01279	R1-R2	0.03979	0.06276	Tidak Bias
		Non-Uniform	F1-F2	0.06276		
6	R-F	0.03929	R1-R2	0.04117	0.04117	Tidak Bias
		Non-Uniform	F1-F2	0.03480		
7	R-F	0.04483	R1-R2	0.01573	0.01573	Bias
		Non-Uniform	F1-F2	0.00935		
8	R-F	0.03894	R1-R2	0.07540	0.07540	Tidak Bias
		Non-Uniform	F1-F2	0.04351		
9	R-F	0.03814	R1-R2	0.03433	0.03717	Bias
		Non-Uniform	F1-F2	0.03717		
10	R-F	0.00419	R1-R2	0.01393	0.03377	Tidak Bias
		Non-Uniform	F1-F2	0.03377		
11	R-F	0.03177	R1-R2	0.02212	0.02212	Bias
		Non-Uniform	F1-F2	0.00366		
12	R-F	0.04552	R1-R2	0.01540	0.03622	Bias
		Non-Uniform	F1-F2	0.03622		
13	R-F	0.06959	R1-R2	0.03150	0.06251	Bias
		Non-Uniform	F1-F2	0.06251		
14	R-F	0.06190	R1-R2	0.03828	0.04694	Bias
		Non-Uniform	F1-F2	0.04694		
15	R-F	0.04789	R1-R2	0.04184	0.04184	Bias
		Non-Uniform	F1-F2	0.03384		
16	R-F	0.05803	R1-R2	0.02237	0.02237	Bias
		Non-Uniform	F1-F2	0.01850		
17	R-F	0.06822	R1-R2	0.10426	0.10426	Tidak Bias
		Non-Uniform	F1-F2	0.04296		
18	R-F	0.02176	R1-R2	0.02061	0.05676	Tidak Bias
		Non-Uniform	F1-F2	0.05676		
19	R-F	0.03434	R1-R2	0.05658	0.05658	Tidak Bias
		Non-Uniform	F1-F2	0.02073		
20	R-F	0.01359	R1-R2	0.01284	0.03586	Tidak Bias
		Non-Uniform	F1-F2	0.03586		
21	R-F	0.04777	R1-R2	0.07401	0.07401	Tidak Bias

Kode Butir	Indeks Kelompok	Indeks Subkelompok	Maksimum	Kesimpulan	
22	R-F	Non-Uniform	F1-F2	0.02063	Bias
		0.05437	R1-R2	0.04870	
23	R-F	Non-Uniform	F1-F2	0.03591	Tidak Bias
		0.02334	R1-R2	0.05363	
24	R-F	Non-Uniform	F1-F2	0.02432	Tidak Bias
		0.02307	R1-R2	0.03896	
25	R-F	Non-Uniform	F1-F2	0.03612	Tidak Bias
		0.01903	R1-R2	0.05286	
26	R-F	Non-Uniform	F1-F2	0.02671	Tidak Bias
		0.05461	R1-R2	0.05770	
27	R-F	Non-Uniform	F1-F2	0.02280	Tidak Bias
		0.02917	R1-R2	0.07254	
28	R-F	Non-Uniform	F1-F2	0.04646	Bias
		0.05285	R1-R2	0.01300	
29	R-F	Non-Uniform	F1-F2	0.02905	Tidak Bias
		0.03948	R1-R2	0.05447	
JUMLAH BIAS BUTIR				12	

2. Metode Khi-Kuadrat Lord

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
1	8.166816621	7.81472	Bias
2	3.251162658	7.81472	Tidak Bias
3	4.0267135	7.81472	Tidak Bias
4	3.895441345	7.81472	Tidak Bias
5	1.001243526	7.81472	Tidak Bias
6	5.124513926	7.81472	Tidak Bias
7	5.120823103	7.81472	Tidak Bias
8	3.860386232	7.81472	Tidak Bias
9	3.599406113	7.81472	Tidak Bias
10	0.167382479	7.81472	Tidak Bias
11	3.454989399	7.81472	Tidak Bias
12	6.291015872	7.81472	Tidak Bias
13	18.24866259	7.81472	Bias
14	11.93467172	7.81472	Bias
15	7.464562276	7.81472	Tidak Bias
16	6.139166726	7.81472	Tidak Bias
17	27.66050698	7.81472	Bias
18	2.189011106	7.81472	Tidak Bias
19	4.483969272	7.81472	Tidak Bias
20	0.56749952	7.81472	Tidak Bias
21	5.523974077	7.81472	Tidak Bias
22	10.45496299	7.81472	Bias

Kode Butir	χ^2_{hitung}	$\chi^2_{tabel} (dk=3, \alpha=0.05)$	Kesimpulan
23	2.713795806	7.81472	Tidak Bias
24	1.865451287	7.81472	Tidak Bias
25	1.592057515	7.81472	Tidak Bias
26	7.139710212	7.81472	Tidak Bias
27	2.940044924	7.81472	Tidak Bias
28	6.557436805	7.81472	Tidak Bias
29	3.125562262	7.81472	Tidak Bias
JUMLAH BIAS BUTIR			5

3. Metode Uji Beda Taraf Sukar (UBTS)

Kode Butir	Δb	$S\Delta b$	d	Kesimpulan
1	-0.39858	0.14876	-2.67930	Bias
2	0.22135	0.24093	0.91872	Tidak Bias
3	-0.17830	0.13019	-1.36949	Tidak Bias
4	0.16060	0.18165	0.88412	Tidak Bias
5	0.14267	0.17371	0.82128	Tidak Bias
6	0.15591	0.07774	2.00558	Bias
7	0.10907	0.11493	0.94901	Tidak Bias
8	0.19069	0.14570	1.30879	Tidak Bias
9	0.25435	2.13452	0.11916	Tidak Bias
10	-0.10726	0.40824	-0.26275	Tidak Bias
11	0.22603	0.18678	1.21014	Tidak Bias
12	-0.00886	0.08219	-0.10786	Tidak Bias
13	-0.10556	0.07010	-1.50587	Tidak Bias
14	-0.29209	0.10336	-2.82610	Bias
15	-0.12841	0.07934	-1.61844	Tidak Bias
16	-0.34168	0.25188	-1.35651	Tidak Bias
17	-0.09384	0.06136	-1.52922	Tidak Bias
18	0.11857	0.17979	0.65949	Tidak Bias
19	0.04683	0.06873	0.68141	Tidak Bias
20	-0.08970	0.12652	-0.70896	Tidak Bias
21	0.09013	0.12868	0.70041	Tidak Bias
22	-0.26227	0.17555	-1.49399	Tidak Bias
23	-0.03078	0.10301	-0.29882	Tidak Bias
24	-0.10119	0.08358	-1.21074	Tidak Bias
25	-0.06750	0.07432	-0.90822	Tidak Bias
26	-0.00771	0.10205	-0.07553	Tidak Bias
27	-0.09921	0.08483	-1.16956	Tidak Bias
28	0.43788	0.32625	1.34216	Tidak Bias
29	0.15886	0.21336	0.74458	Tidak Bias
JUMLAH BIAS BUTIR				3