

ANALISIS BUTIR SOAL MANUAL DAN MENGGUNAKAN *SOFTWARE*

Oleh:

Samsul Hadi

A. Pendahuluan

Supaya soal memenuhi validitas isi penyusunan butir soal diawali dengan mengkaji kompetensi dan indikator yang akan diujikan. Berdasarkan kisi-kisi tersebut kemudian dibuat butir-butir soal. Butir-butir soal yang sudah jadi kemudian diminta untuk ditelaah oleh pihak lain yang dianggap mampu dengan memperhatikan: materi, konstruksi, dan bahasa yang digunakan dalam butir soal.

Telaah materi bertujuan untuk melihat kesesuaian soal dengan kompetensi dan indikator. Telaah konstruksi bertujuan untuk menjamin bahwa soal telah dirumuskan dengan singkat, bebas dari pernyataan yang tidak relevan, bebas dari pernyataan negatif ganda, dan bebas dari pernyataan yang multi interpretasi. Telaah bahasa dilakukan agar soal komunikatif dan sesuai dengan kemampuan peserta tes serta menggunakan bahasa Indonesia yang baku.

Hasil telaah materi, konstruksi, dan bahasa dijadikan masukan perbaikan terhadap butir soal yang ada. Setelah perbaikan dilakukan berdasarkan telaah ketiga hal tersebut, soal siap diujicobakan. Data yang diperoleh dari hasil ujicoba perlu dianalisis untuk mengetahui karakteristik soal atau butir soal secara empiris. Ada dua pendekatan untuk menganalisis data hasil ujicoba soal, yaitu menggunakan teori tes klasik dan menggunakan teori respons butir. Dalam kesempatan ini analisis butir soal hanya akan dilakukan dengan teori tes klasik.

B. Rumus Perhitungan dalam Teori Tes Klasik

Kualitas tes atau butir soal penyusun tes yang baik dapat dilihat dari karakteristiknya. Karakteristik tes atau butir dapat diketahui dengan dua pendekatan teori. Kedua pendekatan tersebut yakni teori tes klasik dan teori respons butir. Teori tes klasik, atau disebut juga teori tes skor murni klasik, didasarkan pada model aditif, yaitu skor amatan merupakan penjumlahan dari

skor sebenarnya dan skor kesalahan pengukuran (Allen & Yen, 1979: 57). Secara matematis pernyataan tersebut dapat dirumuskan sebagai berikut.

$$X = T + E$$

dengan :

X : skor amatan,

T : skor murni,

E : skor kesalahan pengukuran (*error score*).

Kesalahan pengukuran dalam teori tes klasik merupakan kesalahan yang tidak sistematis atau acak. Kesalahan pengukuran merupakan penyimpangan secara teoretis dari skor amatan yang diperoleh dengan skor amatan yang diharapkan. Kesalahan pengukuran yang sistematis dianggap bukan merupakan kesalahan pengukuran.

Asumsi-asumsi yang mendasari teori tes klasik tersebut dijadikan dasar untuk mengembangkan rumus-rumus matematis untuk mengestimasi validitas dan koefisien reliabilitas tes. Validitas dan koefisien reliabilitas pada perangkat tes digunakan untuk menilai kualitas tes. Kualitas tes dalam teori tes klasik juga dapat ditentukan dengan indeks kesukaran dan daya pembeda.

1. Tingkat Kesukaran

Tingkat kesukaran, disimbolkan dengan p , merupakan salah satu parameter butir soal yang sangat berguna dalam analisis soal. Tingkat kesukaran dapat dihitung dengan berbagai cara, yaitu (a) skala kesukaran linear, (b) skala bivariat, (c) indeks Davis, dan (d) proporsi menjawab benar (Bahrul Hayat, dkk., 1999). Secara matematis tingkat kesukaran yang dihitung dengan proporsi menjawab benar dirumuskan dengan:

$$p = \frac{\sum B}{N}$$

dengan keterangan B adalah banyak peserta tes yang menjawab benar, dan N jumlah peserta tes yang menjawab. Dengan rumus tersebut, maka dapat diketahui bahwa jika p mendekati 0, maka soal tersebut terlalu sukar, sedang jika p mendekati 1 maka soal tersebut terlalu mudah. Soal yang terlalu mudah atau

terlalu sukar tidak dapat membedakan kemampuan peserta tes sehingga perlu dibuang.

Menurut Allen dan Yen (1979) tingkat kesukaran butir soal sebaiknya antara 0,3 – 0,7. Pada rentang tersebut informasi tentang kemampuan siswa akan diperoleh secara maksimal. Namun angka tersebut perlu disesuaikan dengan tujuan pengembangan soal. Soal untuk keperluan seleksi, remidi, atau ulangan umum seharusnya mempunyai p yang berbeda-beda untuk mencapai tujuan yang maksimal.

2. Daya Beda

Daya beda merupakan parameter butir soal yang memberikan informasi tentang seberapa besar butir soal tersebut dapat membedakan peserta tes yang skornya tinggi dan peserta tes yang skornya rendah. Daya beda dapat dihitung dengan beberapa cara antara lain dengan menghitung koefisien korelasi *point biserial* dan koefisien korelasi *biserial*. Korelasi *point biserial* secara matematis dirumuskan sebagai berikut.

$$r_{pbis} = \frac{M_p - M_q}{S_t} \sqrt{pq}$$

dimana:

r_{pbis} : koefisien korelasi *point biserial*

M_p : *mean* skor pada tes dari peserta tes yang memiliki jawaban benar pada butir soal

M_q : *mean* skor pada tes dari peserta tes yang memiliki jawaban salah pada butir soal

p : proporsi peserta tes yang menjawab benar pada butir soal

q : $1 - p$

S_t : standar deviasi seluruh skor tes

Hubungan antara korelasi *point biserial* dengan korelasi *biserial* mengikuti rumus sbb:

$$r_{pbis} = r_{bis} \cdot \frac{y}{\sqrt{p \cdot q}}$$

dengan keterangan r_{bis} adalah koefisien korelasi *biserial*, y adalah ordinat p dalam distribusi normal, sedangkan simbol lain sama dengan keterangan sebelumnya. Nilai korelasi *point biserial* selalu lebih rendah dibanding dengan nilai korelasi *biserial*.

3. Efektivitas Distraktor

Soal pilihan ganda perlu memiliki pengecoh, yaitu jawaban yang tidak bernilai benar. Pengecoh perlu dibuat sedemikian rupa sehingga menarik perhatian peserta tes yang belum memiliki konsep yang baik terhadap materi yang diujikan. Allen dan Yen (1979) menyatakan bahwa pengecoh yang baik minimum berindeks 0,1 yang berupa koefisien korelasi *point biserial*, bernilai positif untuk kunci jawaban dan bernilai negatif untuk pengecoh.

4. Kesalahan Pengukuran

Kesalahan Pengukuran (*Standard Error of Measurement, SEM*) membantu penyusun tes dalam memahami kesalahan yang bersifat acak yang mempengaruhi skor peserta tes. Kesalahan pengukuran dihitung dengan rumus sebagai berikut (Bahrul Hayat, dkk., 1999):

$$\sigma_E = \sigma_X \sqrt{1 - \rho_{XX'}}$$

dengan keterangan σ_X adalah standar deviasi dari skor total dan $\rho_{XX'}$ adalah koefisien reliabilitas tes.

5. Reliabilitas Tes

Reliabilitas tes dapat diartikan sebagai keajegan atau konsistensi hasil pengukuran atau hasil tes yang dilakukan pada waktu yang berbeda pada subjek yang sama. Allen dan Yen (1979) menyatakan bahwa tes disebut reliabel jika skor amatan mempunyai korelasi yang tinggi dengan skor yang sebenarnya. Mereka juga menyatakan bahwa reliabilitas merupakan koefisien korelasi antara dua skor amatan yang diperoleh dari hasil pengukuran menggunakan tes yang paralel.

Reliabilitas suatu tes dapat dihitung dengan beberapa rumus atau cara. Rumus atau cara belah dua, alfa (α) Cronbach, Guttman, dan paralel dapat

digunakan. Di antara rumus atau cara tersebut, alfa (α) Cronbach adalah yang paling banyak digunakan karena cocok untuk data benar salah atau angket. Nilai hasil perhitungan dari formula tersebut sering dikatakan sebagai koefisien reliabilitas. Mehrens dan Lehmann (1973) menyatakan bahwa meskipun tidak ada ketentuan umum, tetapi secara luas dapat diterima bahwa untuk tes yang digunakan untuk membuat keputusan secara perorangan harus memiliki koefisien reliabilitas minimal 0,85.

Rumus alfa (α) Cronbach adalah sebagai berikut (Johnson, Penny, & Gordon, 2009):

$$\alpha = \frac{k}{k - 1} \left(1 - \frac{\sum \sigma_i^2}{\sigma_w^2} \right)$$

dengan keterangan k = jumlah butir, σ_i^2 = varians butir, dan σ_w^2 = varians soal (seluruh butir).

C. Analisis Butir Soal dengan ITEMAN

ITEMAN merupakan program analisis instrumen yang membutuhkan file data dalam format ASCII (*text-only*). Semua data yang akan dianalisis harus dientry dalam file tunggal. ITEMAN dapat menganalisis sampai 750 butir sekaligus, dengan jumlah peserta tes hampir tak terbatas.

Format file data dikotomus yang akan dianalisis dengan ITEMAN adalah sebagai berikut:

035 O N 06	Control Line
CDBCADDDCBBCDDBDCABDADDBCBCDBCCBDAB	Key
44	No. Alternatives
YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY	Items to Include
mat1 CDBCABODCBDDDDCACABABDDCAABDACCBC	Examinee #1
mat2 CDBCABDDCBCCDDDCABBBDDBAADDBCCDDB	Examinee #2

Baris pertama disebut *Control Line*. Angka 035 menunjukkan jumlah butir, O menyatakan Omitted, N menyatakan *not-reached item*, dan 06 menunjukkan jumlah karakter identitas peserta tes (maksimal 80 karakter). Kunci jawaban (*key*) maupun jawaban peserta dapat diganti dengan angka. *Items to Include* dapat diganti dengan angka, sedang yang tidak akan di-include harus ditandai dengan N.

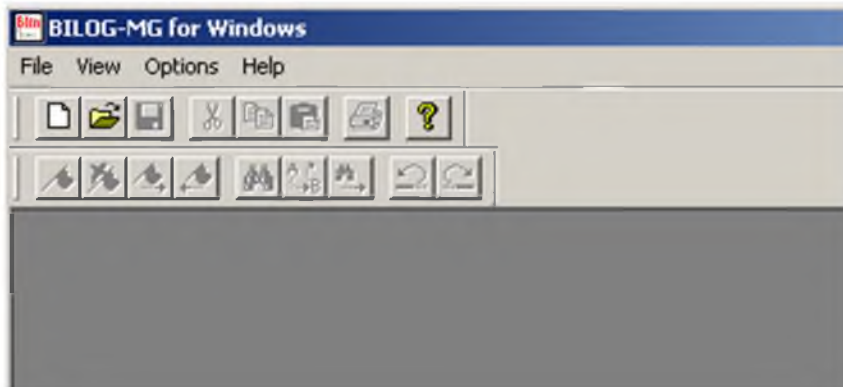
Scale Statistics

Scale:	0

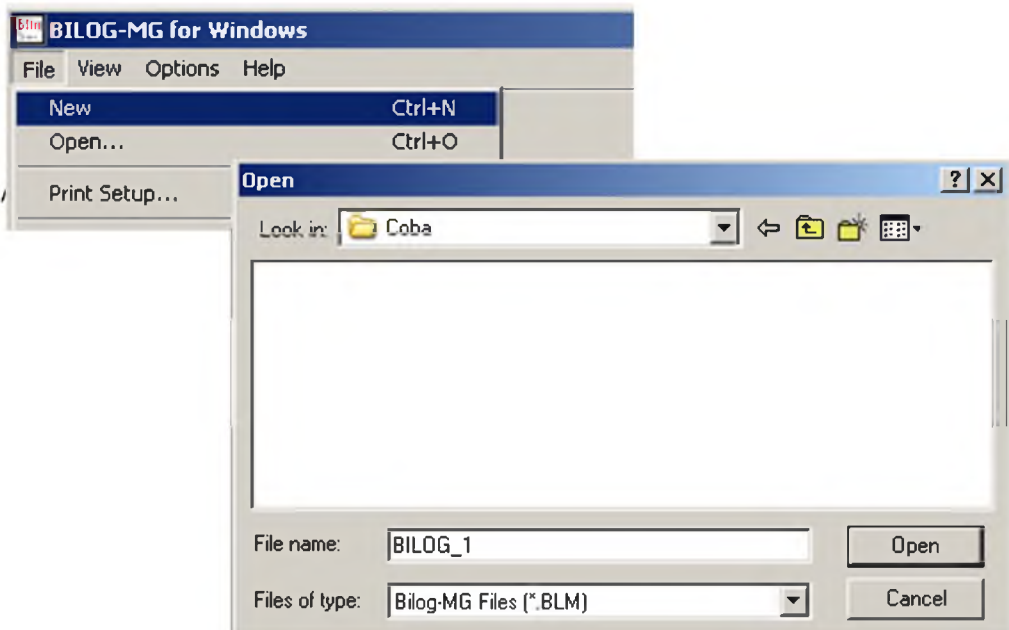
N of Items	24
N of Examinees	26
Mean	14.308
Variance	17.828
Std. Dev.	4.222
Skew	-0.092
Kurtosis	0.580
Minimum	4.000
Maximum	24.000
Median	14.000
Alpha	0.797
SEM	1.900
Mean P	0.596
Mean Item-Tot.	0.421
Mean Biserial	0.604

DAFTAR PUSTAKA

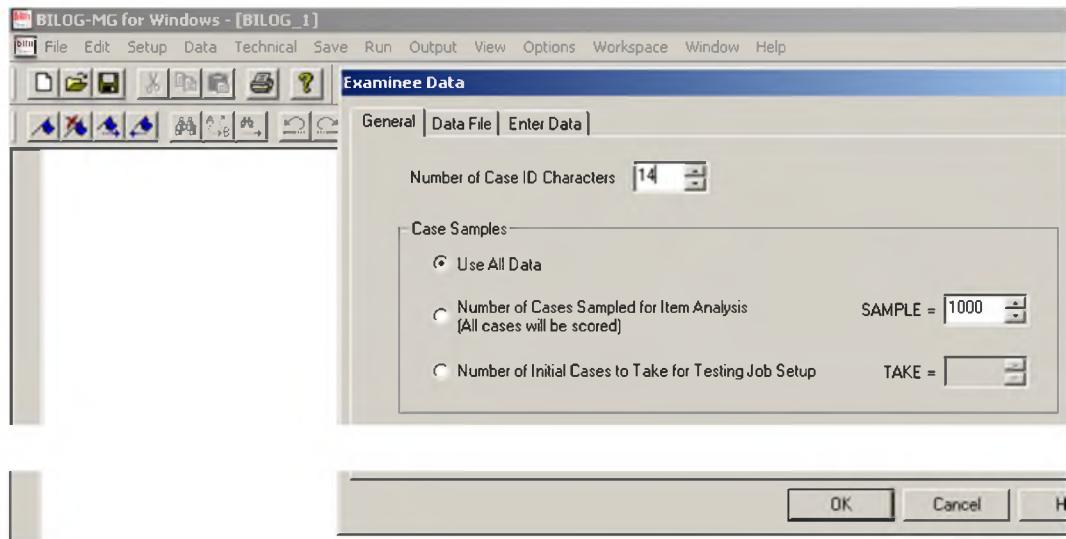
- Allen, M. J & Yen, W. M. (1979). *Introduction to measurement theory*. Belmont: Wadsworth.
- Bahrul Hayat, Sumarno S. Pranata, dan Herwindo Haribowo. (1999). *Manual item and test analysis (ITEMAN)*. Jakarta: Pusbang Sisjian Depdikbud.
- Hulin, C.L., Drasgow, F. & Parsons, C.K. (1983). *Item response theory: Application to psychological measurement*. Homewood, IL: Dow Jones-Irwin.
- Johnson, L.R., Penny, J.A. & Gordon, B. (2009). *Assessing Performance*. New York: The Guilford Press.
- Mehrens, W. A. & Lehmann, I. J. (1973). *Measurement and evaluation in education and psychology*. New York: Hold Rinehart and Wiston.



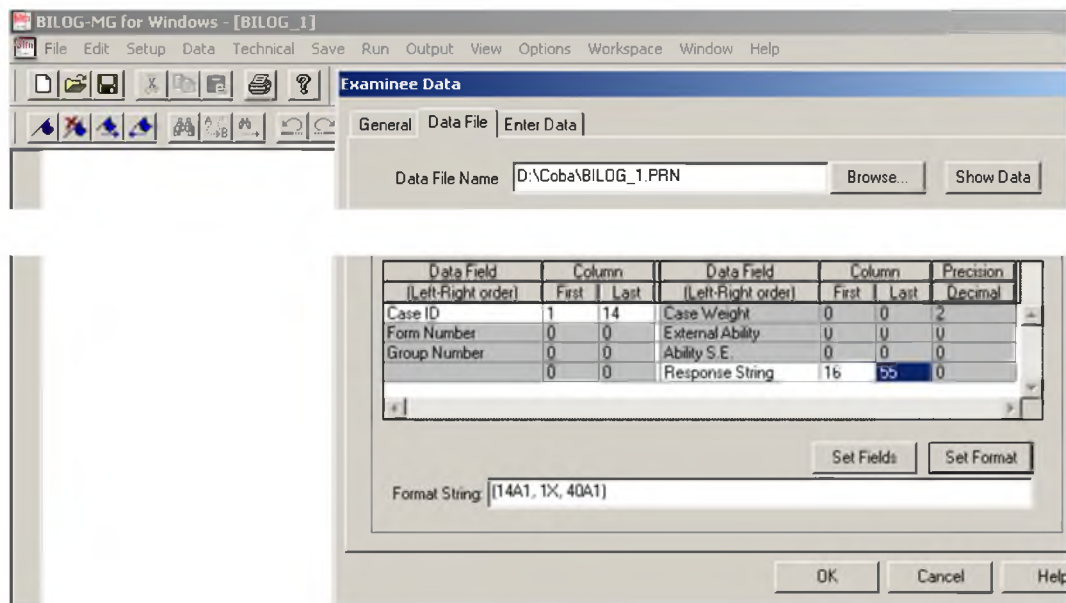
3. Buat sintaks (perintah) untuk menganalisis data yang telah disiapkan dengan cara melihat menu **File** dan memilih **New**, maka akan muncul kotak dialog yang meminta nama folder dan file untuk menyimpan file sintaks. Nama folder ini sebaiknya sama dengan nama folder untuk menyimpan data. Misalkan nama folder-nya diisi Coba dan nama file sintaks yang akan dibuat diisi BILOG_1.BLM. Setelah nama folder dan nama file sintaks diisi kemudian klik **Open**.



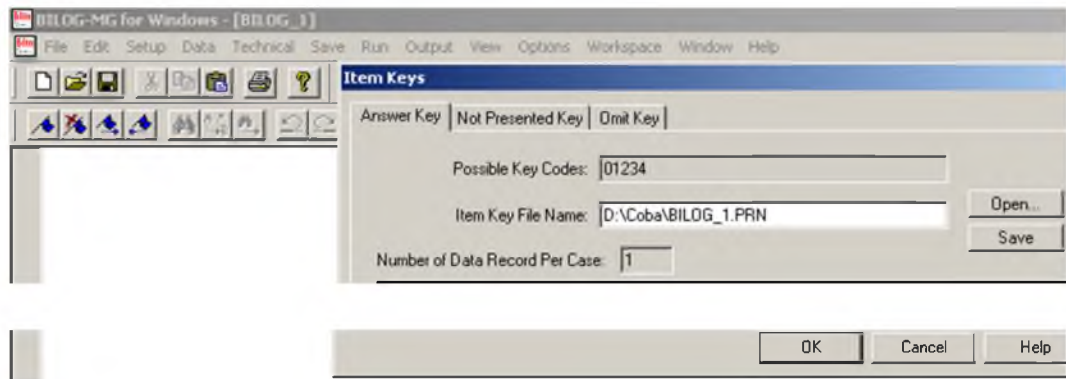
4. Pada tampilan **BILOG-MG** pilih menu **Data** submenu **Examine Data**, isi jumlah karakter identitas peserta sesuai dengan data yang telah disiapkan pada langkah 1.



5. Pada kotak dialog **Examinee Data** klik tab menu **Data File**. Masukkan file data jawaban peserta tes lengkap dengan nama folder-nya. Isi kolom pertama dan terakhir data identitas dan kolom pertama dan terakhir respons jawaban peserta tes, klik tombol **Set Format**, kemudian klik tombol OK.

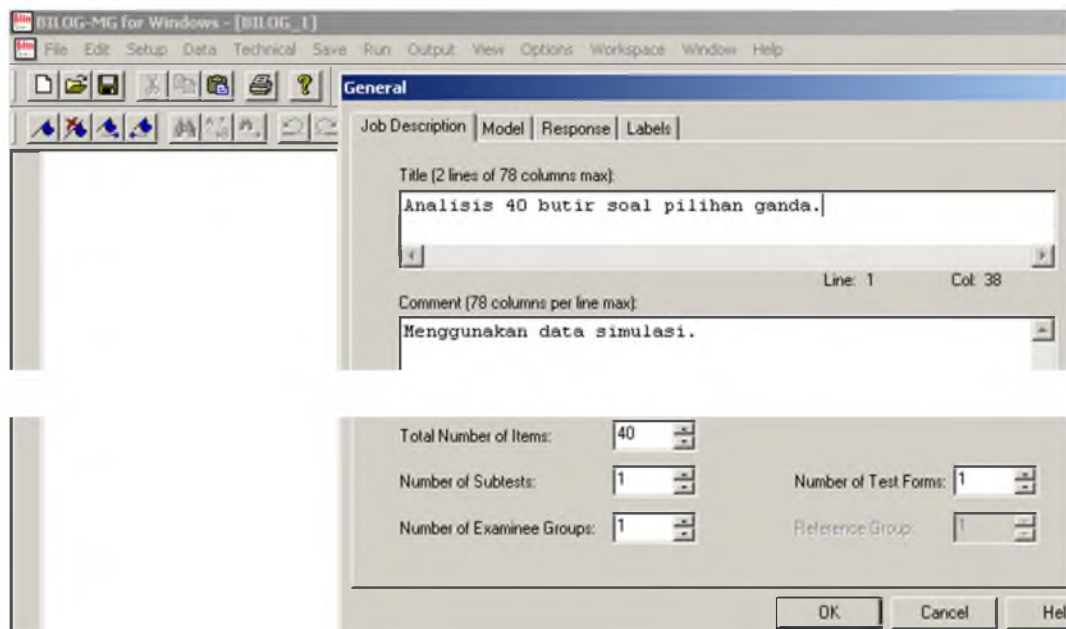


6. Pada tampilan **BILOG-MG** pilih menu **Data** submenu **Item Keys**, isi nama file kunci beserta nama folder-nya pada tempat yang tersedia.

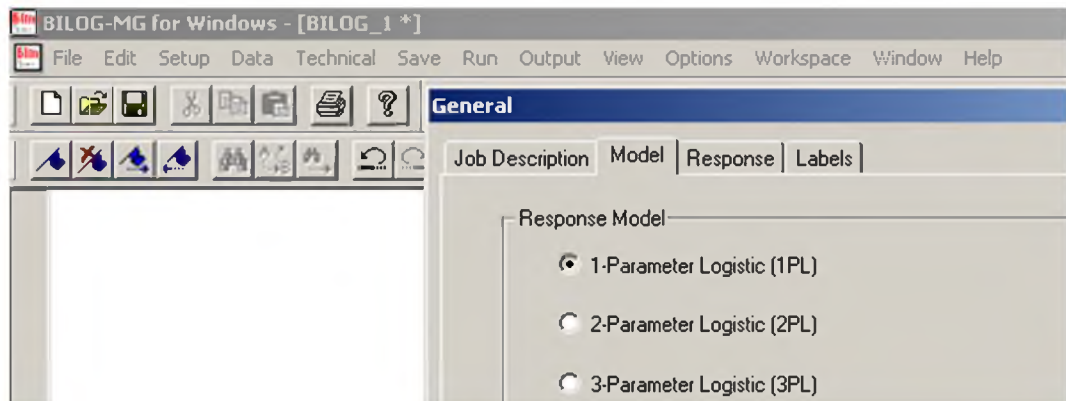


Pada kotak dialog **Item Keys** klik tab menu **Not Presented Key**, isi nama file kunci beserta nama folder-nya. Klik tab menu **Omit Key**, isikan nama file kunci beserta nama folder-nya. Setelah itu klik tombol **OK**. Dalam contoh ini nama file data jawaban peserta tes, kunci jawaban, kode data yang tidak diproses/diabaikan, dan kode data kosong menggunakan nama file yang sama, yaitu BILOG_1.PRN.

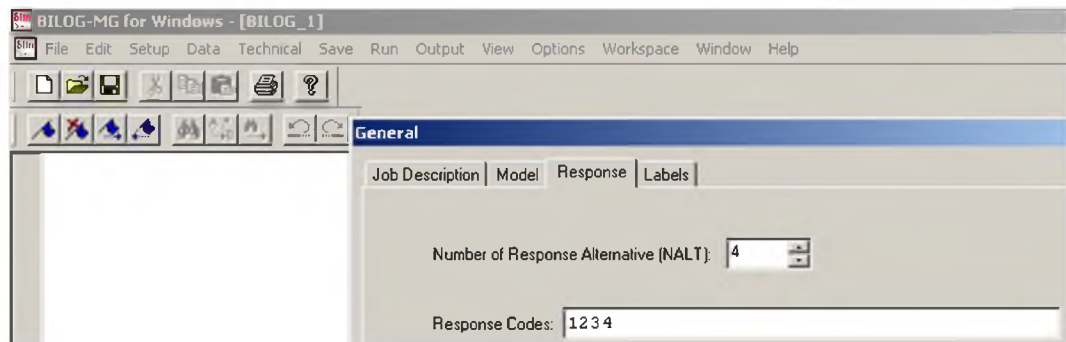
7. Klik menu **Setup** dan pilih submenu **General**. Isi kotak dialog pada **Job Description** dengan judul analisis, komentar, jumlah butir, dan informasi lain yang terkait dengan soal tes. Dalam contoh ini jumlah butir adalah 40.



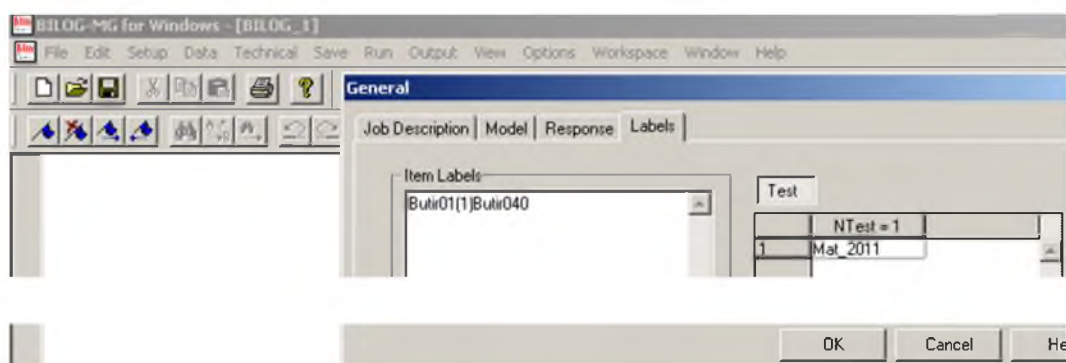
Pada kotak dialog **General** klik tab menu **Model** pada, pilih model IRT yang dikehendaki misalnya 1-Parameter Logistic (1PL), lainnya dapat dibiarkan sesuai default-nya.



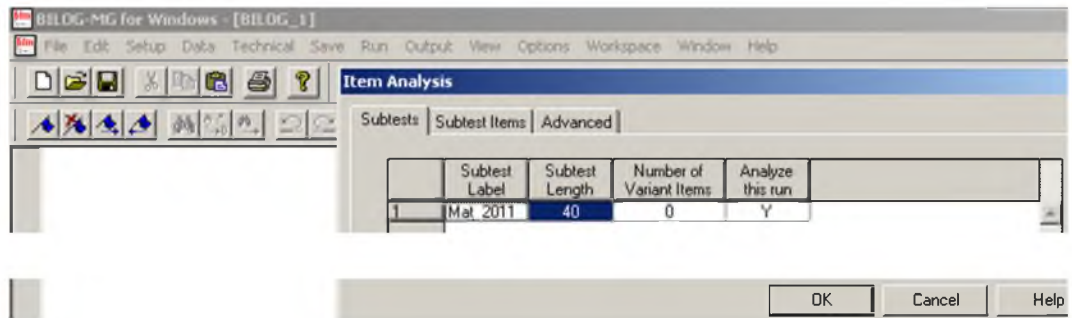
Pada kotak dialog **General** klik tab menu **Response**, isi jumlah alternatif jawaban dan kode jawaban sesuai dengan yang ada dalam file data.



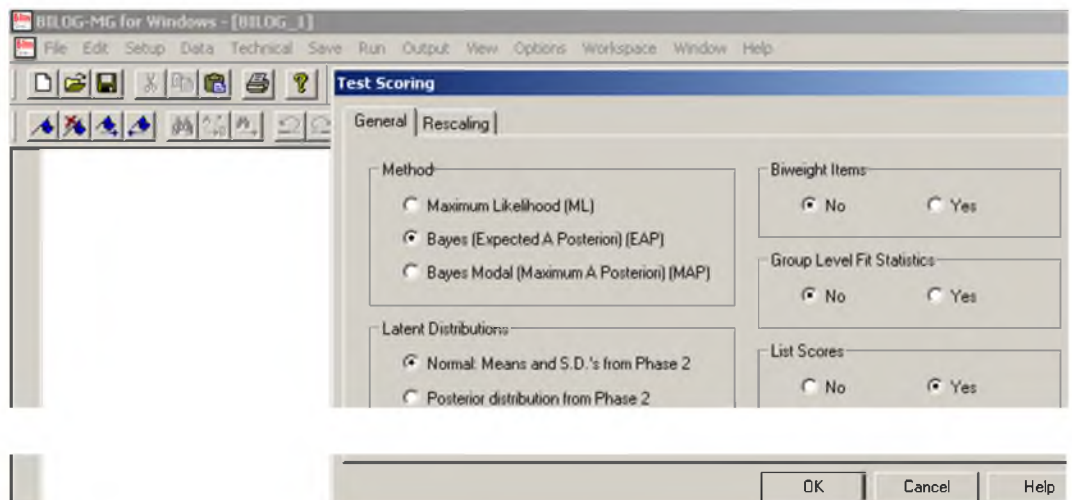
Pada kotak dialog **General** klik tab menu **Label** beri nama butir dan nama tes yang sesuai dengan data yang dianalisis atau biarkan sesuai default-nya, kemudian klik tombol **OK**.



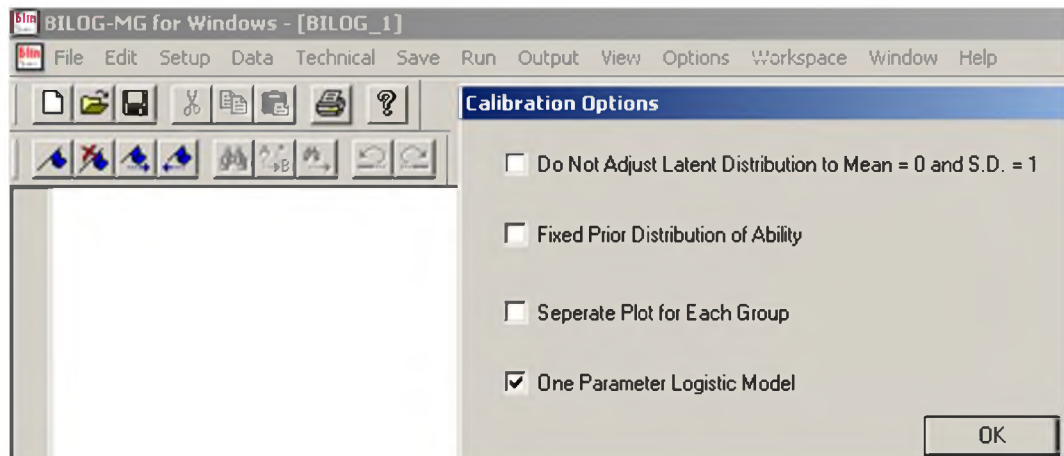
8. Klik menu **Setup**, pilih submenu **Item Analysis**. Pada kotak dialog yang ada, isi menu **Subtest Length** pada tab menu Subtests sebanyak butir test yang ada pada data. Biarkan isi tab menu **Subtest Items** dan **Advanced** pada default-nya, kemudian klik tombol **OK**.



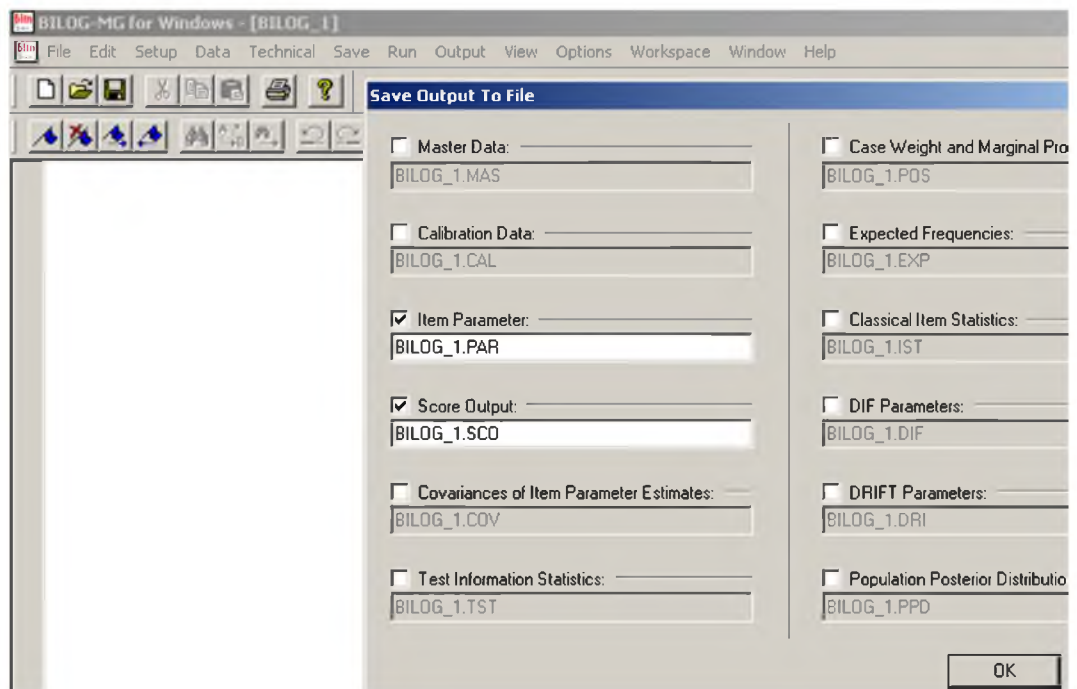
9. Klik menu **Setup** dan pilih submenu **Test Scoring**. Pada kotak dialog yang ada pilih metode estimasi (**Method**) yang ada dalam tab **General** dan reskala hasil estimasi melalui tab **Rescaling** atau biarkan semua pada default-nya, kemudian klik tombol **OK**.



10. Klik menu **Technical** dan submenu **Calibration Options**. Pada kotak dialog yang ada, pilih **One Parameter Logistic Model** jika model yang dipilih adalah 1PL dan akan di-rescale ke model Rasch. Jika langkah ini dilakukan, maka daya beda (a) dari butir yang dianalisis akan diset = 1. Jika model yang dipilih adalah 2PL atau 3PL langkah ini tidak perlu dilakukan.



11. Klik menu **Save** dan pilih output yang akan di-**Save**.



12. Klik menu **Run**, pilih submenu **Build Syntax** untuk membangkitkan sintaks BILOG-MG secara otomatis. Jika hal ini dilakukan akan muncul tampilan sebagai berikut:

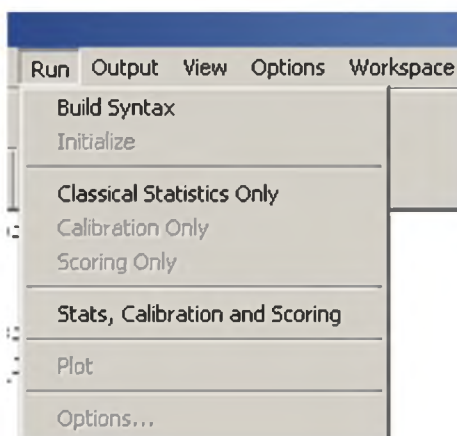
```

Analisis 40 butir soal pilihan ganda.

>COMMENT
Menggunakan data simulasi.
>GLOBAL DFName = 'BILOG_1.PRN',
      NPArm = 1,
      SAVe;
>SAVE PARM = 'BILOG_1.PAR',
      SCORe = 'BILOG_1.SCO';
>LENGTH NITems = (40);
>INPUT NTOtal = 40,
      NALt = 4,
      NIDchar = 14,
      KFName = 'BILOG_1.PRN',
      NFName = 'BILOG_1.PRN',
      OFName = 'BILOG_1.PRN';
>ITEMS INames = (Butir01(1)Butir40);
>TEST1 TName = 'Mat_2011',
      INUmber = (1(1)40);
      (14A1, 1X, 40A1)
>CALIB ACCel = 1.0000,
      RASch;
>SCORE ;

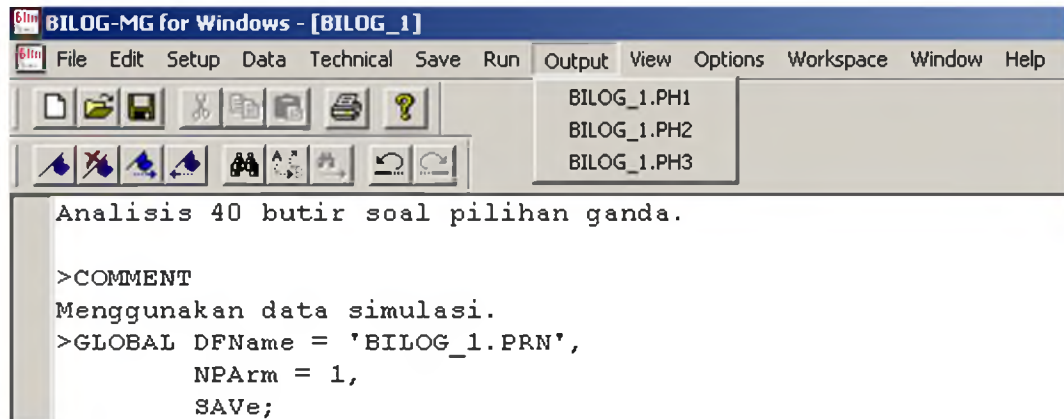
```

Untuk menjalankan sintaks tersebut klik menu **Run** dan pilih submenu **Stats, Calibration and Scoring**.



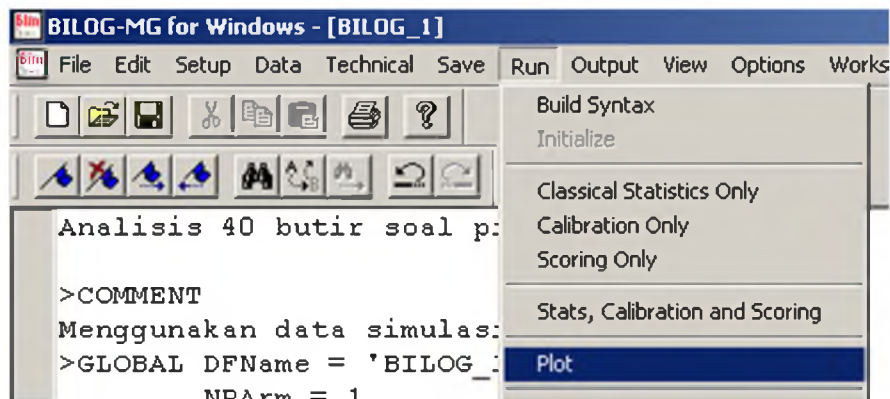
Jika langkah-langkah yang dilakukan benar dan model IRT yang dipilih sesuai dengan kondisi data, maka akan ada tampilan yang menyatakan BILOG-MG telah menganalisis butir soal dengan sukses.

13. Untuk melihat hasil analisis klik menu **Output**

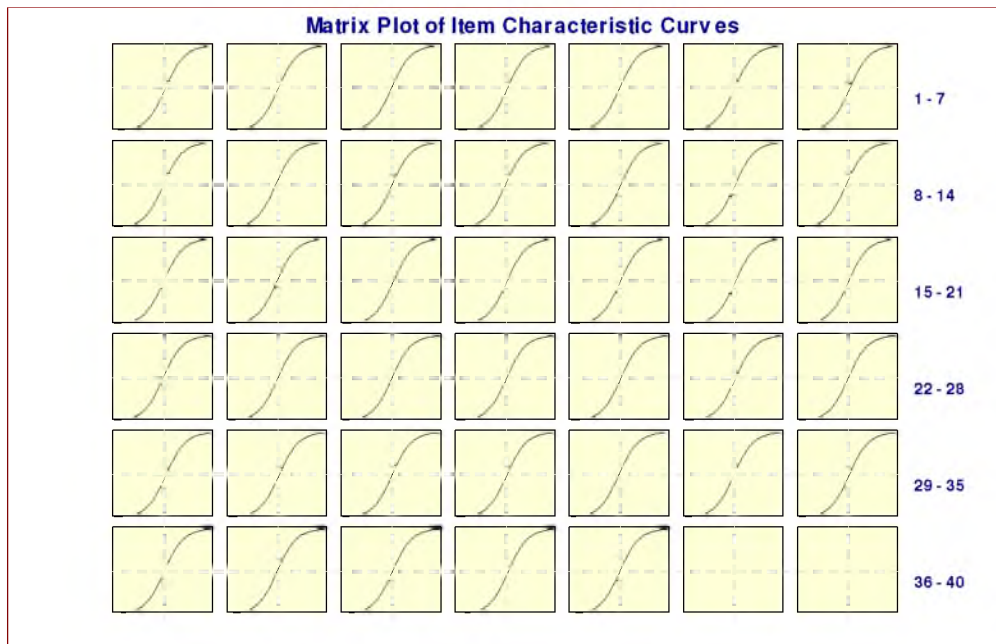


File dengan ekstensi .PH1 merupakan hasil analisis butir menggunakan teori tes klasik oleh BILOG-MG, file dengan ekstensi .PH2 merupakan hasil analisis butir berdasarkan IRT, sedang file dengan ekstensi .PH3 merupakan hasil estimasi kemampuan (θ) peserta tes berdasarkan IRT.

14. Untuk melihat grafik IRT, klik menu **Run** kemudian pilih submenu **Plot**.



Berikut ini merupakan matriks plot ICC dari seluruh butir soal yang dianalisis. Hasil plot dapat di-copy kemudian di-paste di Ms Word. Untuk kembali ke menu BILOG-MG maka **plot harus ditutup**.



Tugas:

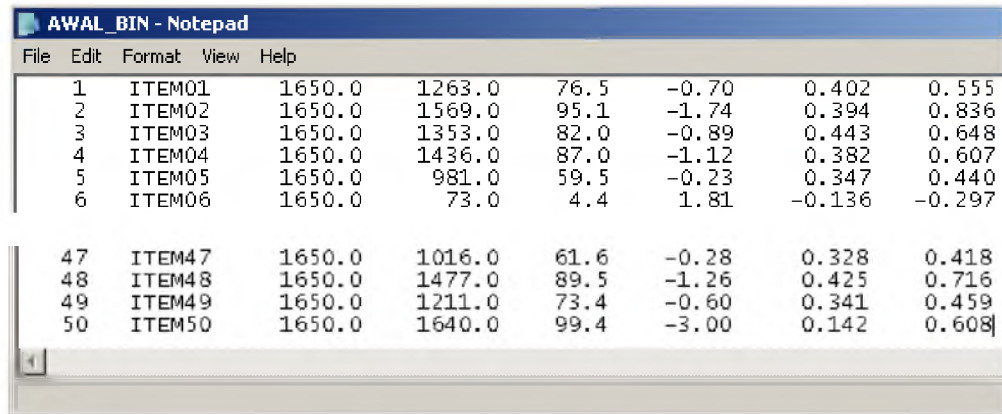
1. Bangkitkan data simulasi UN matematika SMP. Jumlah butir = 40, alternatif jawaban = 4, dan jumlah peserta UN = 500. Analisis data simulasi menggunakan model 1PL. Cermati output-nya. Lihat plotnya.
2. Bangkitkan data simulasi UN IPA SMP. Jumlah butir = 40, alternatif jawaban = 4, dan jumlah peserta UN = 600. Analisis data simulasi menggunakan model 2PL. Cermati output-nya. Lihat plotnya
3. Bangkitkan data simulasi UN Bahasa Indonesia SMP. Jumlah butir = 50, alternatif jawaban = 4, dan jumlah peserta UN = 900. Analisis data simulasi menggunakan model 3PL. Cermati output-nya. Lihat plotnya

MENGOLAH HASIL ANALISIS BILOG-MG DENGAN SPSS

Oleh:
Samsul Hadi

A. Mengolah Output Fase Satu (.PH1)

1. Edit dengan Notepad file output fase 1 hasil analisis BILOG-MG yang telah anda lakukan sehingga tinggal berisi butir soal dan hasil analisis teori tes klasiknya seperti berikut:



Item	Item ID	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	ITEM01	1650.0	1263.0	76.5	-0.70	0.402	0.555		
2	ITEM02	1650.0	1569.0	95.1	-1.74	0.394	0.836		
3	ITEM03	1650.0	1353.0	82.0	-0.89	0.443	0.648		
4	ITEM04	1650.0	1436.0	87.0	-1.12	0.382	0.607		
5	ITEM05	1650.0	981.0	59.5	-0.23	0.347	0.440		
6	ITEM06	1650.0	73.0	4.4	1.81	-0.136	-0.297		
47	ITEM47	1650.0	1016.0	61.6	-0.28	0.328	0.418		
48	ITEM48	1650.0	1477.0	89.5	-1.26	0.425	0.716		
49	ITEM49	1650.0	1211.0	73.4	-0.60	0.341	0.459		
50	ITEM50	1650.0	1640.0	99.4	-3.00	0.142	0.608		

Simpan file output fase 1 tersebut dengan nama baru, misalnya AWAL_BIN.TXT (BIN bisa diganti sesuai nama mata pelajaran).

2. Buka SPSS, buat sintaks untuk membaca output fase 1 yang sudah diedit melalui menu File – New – Syntax. Ketik sintaks sebagai berikut (sesuaikan nama folder dan file yang akan dibaca SPSS dengan data saudara):

```
DATA LIST FILE = 'D:\Coba\AWAL_BIN.TXT'  
/BUTIR 9-14(A) TRIED 20-25 RIGHT 30-35 PERCENT 39-43 PEARSON_COR 57-62 BISER_COR 66-71.  
FORMAT TRIED RIGHT(F6.1) PERCENT (F5.1) PEARSON_COR BISER_COR (F6.3).  
EXECUTE.
```

Jalankan sintaks tersebut melalui menu Run All pada menu Syntax Editor SPSS. Cek data SPSS yang muncul dengan output fase 1 BILOG-MG. Jika ada perbedaan sesuaikan kolom untuk tiap variabel pada sintaks di atas.

3. Identifikasi butir yang valid dengan ketentuan daya beda minimal 0,3 menggunakan menu Transform – Recode Into Different Variables dari variabel BISER_COR menjadi variabel Keterangan. Sintaks yang dihasilkan adalah sbb:

```
STRING Keterangan (A12) .  
RECODE  
  BISER_COR  
  (0.3 thru Highest='Valid') (ELSE='Tidak valid') INTO Keterangan .  
EXECUTE .
```

Select sintaks tersebut kemudian alankan sintaks tersebut melalui menu Run Selection pada menu Syntax Editor SPSS.

Tugas A:

- Lakukan analisis deskriptif variabel TRIED, RIGHT, PERCENT, PEARSON_COR, BISER_COR, dan Keterangan yang terbentuk oleh sintaks SPSS di atas.

B. Mengolah Output Fase 2 (.PH2)

1. Edit dengan Notepad file fase 2 hasil analisis BILOG-MG yang telah anda lakukan sehingga tinggal berisi butir soal dan hasil analisis teori respons butir dan yakinkan informasi setiap butir direkam dalam 3 baris (termasuk baris kosong) seperti sbb:

Item	Mean (SD)	Slope (SE)	Threshold (SE)	ASYP (SE)	CHISQ (SE)	PROB (SE)	Other
ITEM01	0.832 0.038*	0.582 0.006*	-1.430 0.066*	0.503 0.005*	0.000 0.000*	5.3 (0.8028)	9.0
ITEM02	2.013 0.074*	0.582 0.006*	-3.458 0.126*	0.503 0.005*	0.000 0.000*	31.2 (0.0000)	5.0
ITEM03	1.063 0.043*	0.582 0.006*	-1.825 0.073*	0.503 0.005*	0.000 0.000*	29.3 (0.0003)	8.0

ITEM48	1.485 0.053*	0.582 0.006*	-2.551 0.091*	0.503 0.005*	0.000 0.000*	29.6 (0.0001)	7.0
ITEM49	0.715 0.036*	0.582 0.006*	-1.228 0.061*	0.503 0.005*	0.000 0.000*	6.6 (0.5828)	8.0
ITEM50	3.313 0.190*	0.582 0.006*	-5.689 0.326*	0.503 0.005*	0.000 0.000*	3.3 (0.3505)	3.0

Simpan file output fase tersebut dengan nama baru, misalnya PH2_BIN.TXT (BIN bisa diganti sesuai nama mata pelajaran).

4. Buka SPSS, buat sintaks untuk membaca output fase 2 yang sudah diedit melalui menu File – New – Syntax. Ketik sintaks sebagai berikut (sesuaikan nama folder dan file dengan data saudara):

```
DATA LIST FILE='D:\Coba\PH2_BIN.TXT' RECORDS =3
/1 ITEM 2-9 (A) SLOPE 24-30 THRESHOLD 35-41 ASYMP TOTE 57-63 CHISQ 68-74
/2 PROB 69-74.
FORMAT SLOPE THRESHOLD ASYMP TOTE (F7.3) CHISQ (F7.1) PROB (F7.4).
EXECUTE.
```

Jalankan sintaks tersebut melalui menu Run All pada menu Syntax Editor SPSS. Cek data SPSS yang muncul dengan output fase 2 BILOG-MG. Jika ada perbedaan sesuaikan kolom untuk tiap variabel pada sintaks di atas.

5. Menggunakan sintaks SPSS, ubah variabel SLOPE, THRESHOLD, ASYMP TOTE. dan CHISQ masing-masing menjadi variabel Keterangan_a, Keterangan_b, Keterangan_c, dan Keterangan_Fit dengan ketentuan sbb:

Parameter/PROB	Nilai	Keterangan
a	0.4 s/d 2	a baik
b	-2 s/d 2	b baik
c	0 s/d 1/jml alternatif	c baik
PROB	> 0,05	Fit model

Sintaks yang harus dibuat kira-kira sebagai berikut:

```

STRING Keterangan_a (A15) .
RECODE
  SLOPE
  (.4 thru 2='a baik') (ELSE='a tidak baik') INTO Keterangan_a .
STRING Keterangan_b (A15) .
RECODE
  THRESHOLD
  (-2 thru 2='b baik') (ELSE='b tidak baik') INTO Keterangan_b .
STRING Keterangan_c (A15) .
RECODE
  ASYMPTOTE
  (0 thru .25='c baik') (ELSE='c tidak baik') INTO Keterangan_c .
STRING Keterangan_Fit (A15) .
RECODE
  PROB
  (0.05 thru Highest='Fit model') (ELSE='Tidak fit model') INTO
  Keterangan_Fit .
EXECUTE .

```

Tugas B:

- Lakukan analisis deskriptif terhadap variabel Keterangan_a, Keterangan_b, Keterangan_c, dan Keterangan_Fit (seduaikan dengan model IRT yang anda gunakan, jika 1 PL variabel yang dianalisis hanya Keterangan_b dan Keterangan_Fit, jika 2 PL variabelnya Karaeangan_a, Keterangan_b, dan keterangan_Fit, jika 3 PL, variabelnya semua).
- Jika anda telah menganalisis data anda untuk model 1 PL, 2 PL, dan 3 PL bandingkan model mana yang paling banyak menghasilkan butir yang Fit model. Jika anda belum menganalisis data anda untuk model 1 PL, 2 PL, dan 3 PL lakukanlah, kemudian bandingkan model mana yang paling banyak menghasilkan butir yang Fit model.
- Model IRT mana yang paling cocok untuk menganalisis data anda?

C. Mengolah Output Fase 3 (.PH3)

1. Edit dengan Notepad file fase 3 hasil analisis butir yang telah anda lakukan sehingga tinggal ID_TESTEE, TRIED, RIGHT, PERCENT, ABILITY, SE, dan MARGINAL PROB dari semua peserta tes. Yakinkan bahwa setiap Record ditulis dalam 2 baris. Dalam contoh berikut ID_TESTEE pertama adalah 10001, sedangkan ID_TESTEE terakhir adalah 11650.

MAT_OK_1_PH3 - Notepad								
File Edit Format View Help								
1	10001							
1.00	MAT_09	39	13	33.33	-1.3332	0.0174	0.000000	
1	10002							
1.00	MAT_09	39	11	28.21	-1.3359	0.0478	0.000000	
1	10003							
1.00	MAT_09	39	27	69.23	-0.4389	0.0701	0.000000	
1	11647							
1.00	MAT_09	39	12	30.77	-1.3339	0.0233	0.000000	
1	11648							
1.00	MAT_09	39	22	56.41	-0.4488	0.0622	0.000000	
1	11649							
1.00	MAT_09	39	30	76.92	-0.1139	0.4296	0.000000	
1	11650							
1.00	MAT_09	39	19	48.72	-0.7262	0.4136	0.000000	

- Buka SPSS, buat sintaks untuk membaca output fase 3 yang sudah diedit melalui menu File – New – Syntax. Ketik sintaks sebagai berikut (sesuaikan nama folder dan file dengan data saudara):

```
DATA LIST FILE ='D:\Coba\MAT_OK_1_PH3.TXT ' RECORDS = 2
/1 GROUP 3-4 ID_TESTEE 7-36 (A)
/2 TRIED 23-25 RIGHT 30-32 PERCENT 36-41 ABILITY 46-53 SE 57-64.
FORMAT TRIED RIGHT (F3.0) PERCENT (F11.2) ABILITY SE (F8.4).
EXECUTE.
```

Jalankan sintaks tersebut melalui menu Run All pada menu Syntax Editor SPSS. Cek data SPSS yang muncul dengan output fase 3 BILOG-MG. Jika ada perbedaan sesuaikan kolom untuk tiap variabel pada sintaks di atas.

Pada SPSS Data Editor, variabel ID_TESTEE berarti identitas peserta tes, TRIED adalah jumlah usaha yang telah dilakukan peserta tes untuk menjawab butir soal yang tersedia, RIGHT menunjukkan jumlah jawaban benar dari peserta tes, PERCENT adalah persentase jawaban benar yang identik dengan NILAI skala 100 pada penilaian secara klasik. ABILITY merupakan hasil estimasi IRT atas kemampuan peserta tes dengan kisaran angka antara -4 sampai +4 sesuai dengan rentang skor baku. ABILITY hasil estimasi IRT ini kadang juga diubah dalam skala 100 dengan rumus $ABILITY_{100} = 12,5 * ABILITY + 50$.

Tugas C:

- Buat variabel NILAI_100 dan ABILITY_100 berdasarkan penjelasan di atas.
- Lakukan analisis deskriptif yang mencakup Mean, Maximum, Minimum, Standard Deviation dari variabel NILAI_100, ABILITY, ABILITY_100 dari data yang diperoleh dari sintaks SPSS di atas. Cermati hasilnya.
- Buat diagram pencar (scatter plot) NILAI_100 klasik dengan ABILITY dan ABILITY_100.
- Bagaimana model matematik hubungan antara NILAI_100 klasik dengan ABILITY dan ABILITY_100?