



**PENGANTAR AUTHENTIC ASSESSMENT
(Penilaian Otentik)**

**Penerapan Sistem Penilaian
Pada Kurikulum Berbasis Kompetensi**

**Oleh
Ngadimun Hd**

Disampaikan dalam Pelatihan Sistem Pengujian bagi
Dosen Unila di Bandar Lampung Tgl. 20 s.d. 31 Desember 2004

**BANDAR LAMPUNG
2004**

ANALISIS BUTIR SOAL DENGAN KOMPUTER DAN MENAFSIRKANNYA¹

Oleh: Ngadimun Hd.²

Pendahuluan

Micro Computer Adaptive Test (MicroCat) menyediakan program *Item analisis (Iteman)* versi 3.50 yaitu program analisis butir soal dengan komputer, yang edisi keempatnya dipublikasikan mulai tahun 1988. Program ini sangat sederhana, jika Anda mau mendalami sedikit saja, akan dapat menguasainya. Jika masih juga ada kesulitan, nara sumber dengan senang hati mau memberikan penjelasan kembali. Pekerjaan analisis butir soal merupakan salah satu jenis kegiatan dosen/guru yang sangat bermanfaat dalam menafsirkan hasil belajar siswa dan untuk mengetahui kualitas soal, tetapi sayangnya masih banyak dosen/guru belum dapat melakukan.

Langkah-langkah kerja Program Iteman

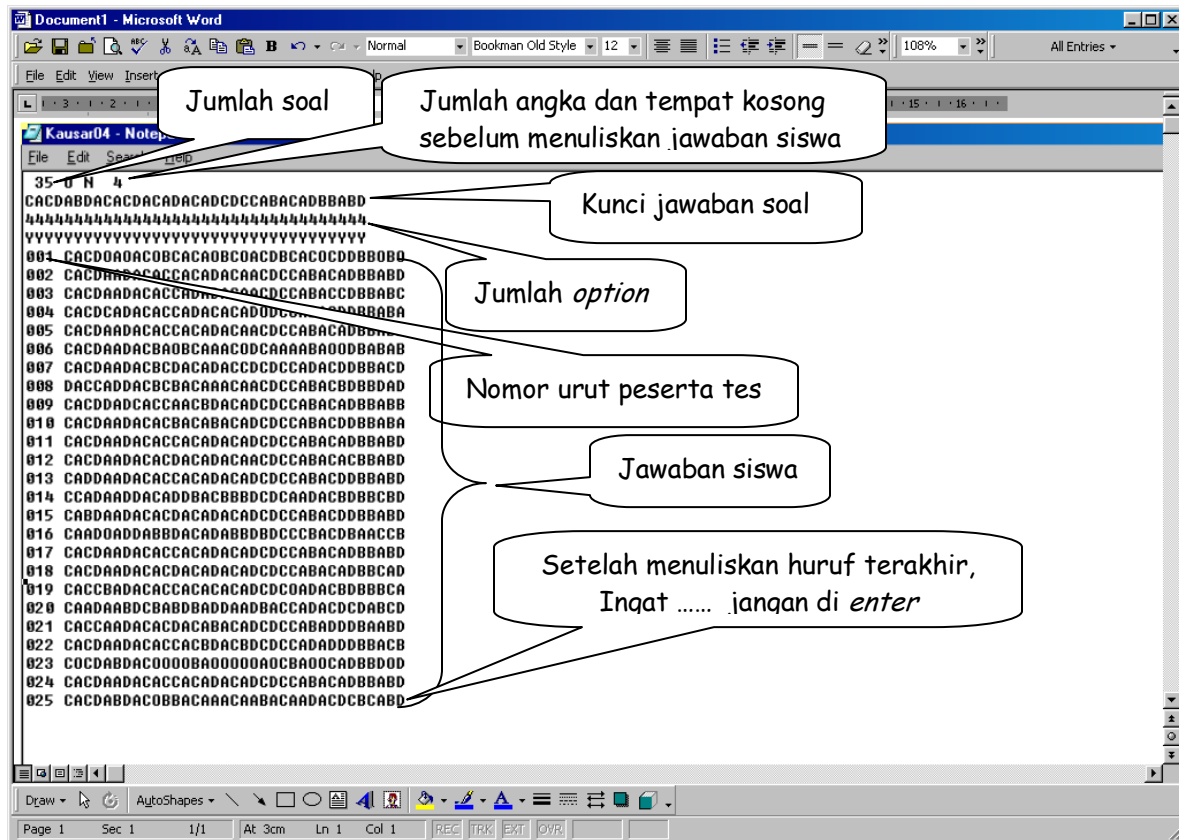
Program ini hanya dapat menganalisis butir soal ragam pilihan ganda. Analisis butir soal dengan program ini relatif mudah, yang penting diperlukan ketelitian dalam memasukkan data (*data entry*), karena kesalahan memasukkan data akan berakibat tidak tepat hasil analisisnya. Langkah-langkah untuk *entry* data (dengan file baru) adalah sbb.

1. klik *Star*
2. pilih *Program*
3. pilih *Accessories*
4. pilih dan klik *Notepad*
5. simpan dengan klik *file*
6. pilih dan klik *Save as*, lalu tulis nama file data, misalnya : *biologi2* (jangan lupa, paling banyak 8 huruf/angka)
7. *entry* data dimulai, lebih cepat jika dilakukan oleh dua orang, seorang membaca jawaban siswa dan seorang menuliskan. Agar tidak keliru antara jawaban b dan d, maka pembacaannya: a untuk a, be untuk b, ce untuk c, *del* untuk d, dan e untuk e.
8. Agar data tidak hilang saat listrik mati, maka sebentar-sebentar disimpan dengan klik *File* dan klik *Save*.
9. Tampilan file data seperti terlihat pada halaman berikut.

Contoh tampilan file data dengan jumlah soal = 35 dan option = 4

¹ Disampaikan dalam Pelatihan Sistem Pengujian bagi Dosen Unila di B. Lampung Tgl.20-31 Desember 2004.

² Dosen FKIP Unila.



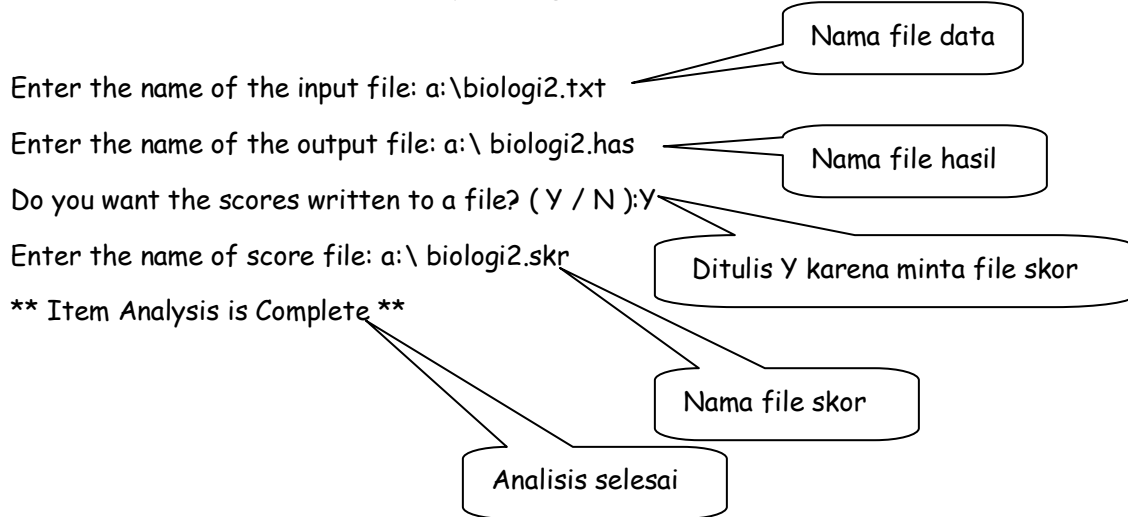
Gambar 1. Tampilan file data dengan Notepad pada Windows

Analisis Data

1. Buka Program Iteman, dengan klik Star, pilih Program, pilih dan klik Iteman.
2. Isi nama file data, pada *Enter the name of the input file:* a:\biologi2.txt lalu tekan enter.
3. Isikan nama file hasil, pada *Enter the name of the output file:* a:\ biologi2.has lalu tekan enter.
4. Lalu muncul pertanyaan *Do you want the scores written to a file? (Y / N)*. maka tulislah Y dan tekan enter.
5. Isikan nama file skor, pada *Enter the name of the score file:* a:\ biologi2.skr Lalu tekan enter, selesai.
6. Tampilannya seperti terlihat pada gambar di halaman berikut.

Tampilan Perintah Analisis Data

MicroCat (tm) Testing System
Copyright © 1982, 1984, 1986, 1988 by Assessment Systems Corporation
Beta-Test version - Univ. of Pittsburgh
Item and Test Analysis Program -- ITEMAN (tm) Version 3.00



Gambar 2. Tampilan analisis butir soal dengan Iteaman

Hasil Analisis Butir Soal

Buka hasil analisis pada program MsWord dengan cara:

1. klik Star, pilih Program, pilih dan klik Microsoft Word
2. klik File, klik Open, dan cari file hasil pada drive D
3. Contoh file hasil analisis seperti pada halaman berikut.

Contoh Tampilan File Hasil Analisis

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics								
		Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low	High	Point Biser.	Key			
1	0-1	.96	.00	.05	A	.00	.00	.00					
					B	.00	.00	.00					
					C	.96	1.00	1.00	.05	*			
					D	.04	.00	.00	-.05				
					Other	.00	.00	.00					
2	0-2	.92	.29	.46	A	.92	.71	1.00	.46	*			
					B	.00	.00	.00					
					C	.04	.14	.00	-.30				
					D	.00	.00	.00					
					Other	.04	.00	.00	-.33				
6	0-6	.08	-.29	-.33	A	.88	.71	1.00	.31	?			
					B	.08	.29	.00	-.33	*			
					CHECK THE KEY				C	.00	.00	.00	
					B was specified, A works better				D	.04	.00	.00	-.05
									Other	.00	.00	.00	

Dan seterusnya

Pada halaman berikut, ditampilkan resume hasil analisis dan di sebelah kanannya skor yang diperoleh setiap peserta ujian.

There were 25 examinees in the data file.

Scale Statistics

N of Items 35
N of Examinees 25
Mean 26.520
Variance 33.930
Std. Dev. 5.825
Skew -0.601
Kurtosis -1.042
Minimum 15.000
Maximum 33.000
Median 29.000
Alpha 0.874
SEM 2.069
Mean P 0.758
Mean Item-Tot. 0.422
Mean Biserial 0.614
Max Score (Low) 23
N (Low Group) 7
Min Score (High) 32
N (High Group) 8

Scores for examinees from file D:\KAUTSAR.TXT	
001	19.00
002	32.00
003	29.00
004	26.00
005	32.00
006	19.00
007	29.00
008	25.00
009	27.00
010	30.00
011	33.00
012	32.00
013	31.00
014	18.00
015	32.00
016	17.00
017	33.00
018	32.00
019	24.00

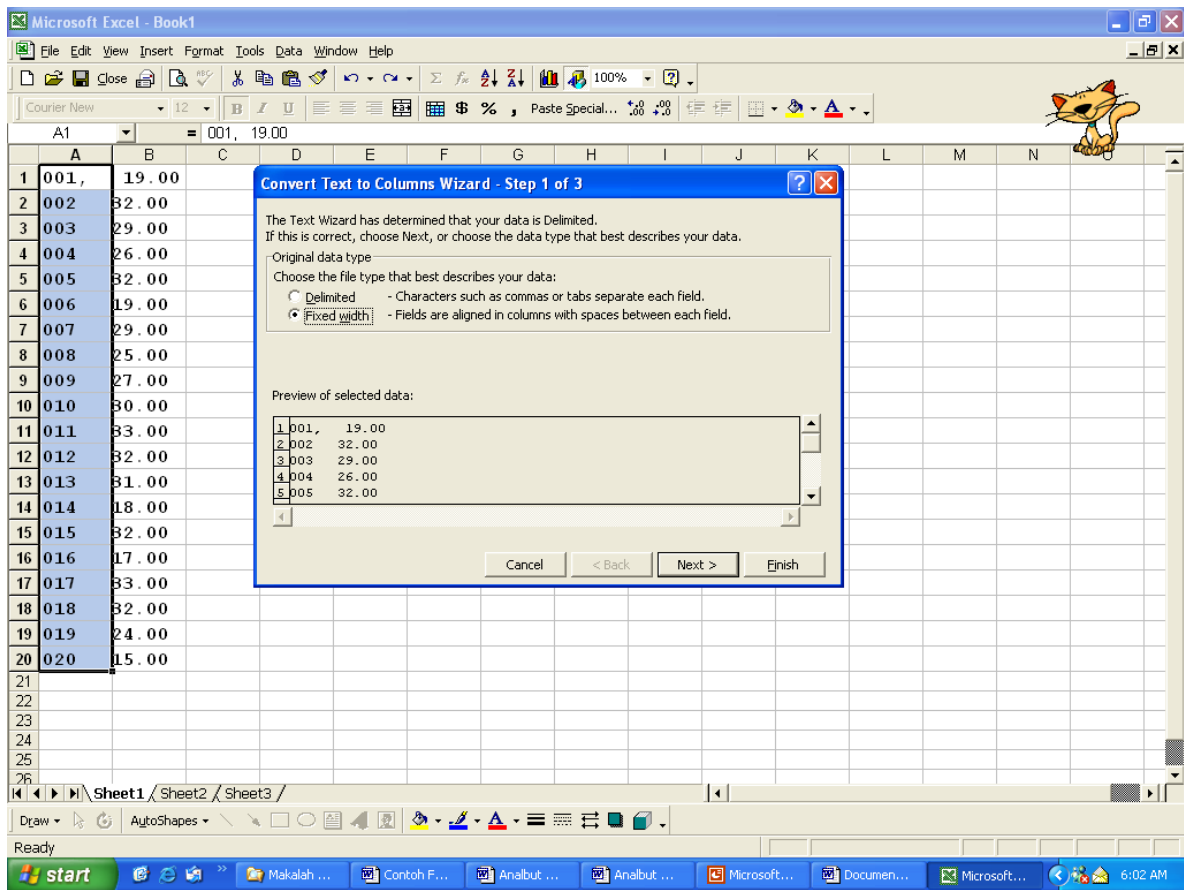
Skor siswa no.urut
001 s.d. 020

Untuk menghitung nilai dengan skala 10 maka skor 20 orang peserta tes di atas, kita transfer ke program Excel, dengan cara:

- Buat tanda koma di antara nomor peserta dg skor (satu peserta saja)
- Copy dan paste di Excel
- Klik data
- pilih text to colum
- klik fixed width
- Klik next, dan klik finish
- Lalu hitung nilai, misal: membagi skor dengan jumlah soal lalu dikalikan 10
- Nilai akan dimasukkan pada cel C1, maka formula penghitungannya adalah:

=B1/35*10 lalu Enter, maka akan diperoleh nilai 5,4.

Ini membacanya, skor peserta tes 001 yang terletak pada cel B1 (kolom B baris pertama), dibagi skor maksimum (jumlah soal 35) lalu dikalikan 10.



Skor dari Item Masih satu kolom		Sudah jadi 2 kolom		
No. Urut	Skor	No. Urut	Skor	Nilai
001,	19.00	001,	19	5.43
002	32.00	2	32	9.14
003	29.00	3	29	8.29
004	26.00	4	26	7.43
005	32.00	5	32	9.14
006	19.00	6	19	5.43
007	29.00	7	29	8.29
008	25.00	8	25	7.14
009	27.00	9	27	7.71
010	30.00	10	30	8.57
011	33.00	11	33	9.43
012	32.00	12	32	9.14
013	31.00	13	31	8.86
014	18.00	14	18	5.14
015	32.00	15	32	9.14
016	17.00	16	17	4.86
017	33.00	17	33	9.43
018	32.00	18	32	9.14
019	24.00	19	24	6.86
020	15.00	20	15	4.29
021	29.00	21	29	8.29
022	26.00	22	26	7.43
023	17.00	23	17	4.86
024	33.00	24	33	9.43
025	23.00	25	23	6.57

3. Menafsirkan hasil analisis butir soal

Untuk menafsirkan hasil analisis soal diperlukan suatu kriteria yang digunakan sebagai patokannya. Beberapa tokoh evaluasi, Sumadi Suryabrata, (1987); Dali S. Naga, (1992); Suharsimi Arikunto, (1993); dan Saifuddin Azwar, (1996); telah menetapkan kriteria dimaksud, yang satu dengan lainnya relatif sama, yaitu seperti pada tabel di halaman berikut.

Tabel 1. Kriteria kualitas butir soal

Kriteria	Indeks	Klasifikasi
<i>Prop Corect</i> (Tingkat kesukaran) (p)	0,000 - 0,250	Sukar
	0,251 - 0,750	Sedang
	0,751 - 1,000	Mudah
<i>Biser</i> (Daya beda) (D)	$D \leq 0,199$	Sangat Rendah
	0,200 - 0,299	Rendah
	0,300 - 0,399	Sedang
	$D \geq 0,400$	Tinggi
Proporsi jawaban <i>Prop Endorsing</i>	0,000 - 0,010	Kurang
	0,011 - 0,050	Cukup
	0,051 - 1,000	Baik
Reliabilitas (Alpha)	0,000 - 0,400	Rendah
	0,401 - 0,700	Sedang
	0,701 - 1,000	Tinggi

Untuk memudahkan memilah-milah butir soal mana yang perlu direvisi atau didrop dianjurkan untuk menggunakan kriteria seperti pada tabel berikut

Tabel 2. Kriteria kualitas soal untuk kepentingan pemilahan butir

Kriteria	Indeks	Klasifikasi	Penafsiran
Tingkat kesukaran (p)	0,000 - 0,099	Sangat Sukar	Dibuang / perlu revisi total
	0,100 - 0,299	Sukar	Perlu direvisi
	0,300 - 0,700	Sedang	Baik
	0,701 - 0,900	Mudah	Perlu direvisi
	0,901 - 1,000	Sangat Mudah	Dibuang / perlu direvisi total
Daya beda (D)	$D \leq 0,199$	Sangat Rendah	Dibuang / perlu direvisi total
	0,200 - 0,299	Rendah	Perlu direvisi
	0,300 - 0,399	Sedang	Sedikit atau tanpa revisi
	$D \geq 0,400$	Tinggi	Bagus Sekali
Proporsi jawaban	0,000 - 0,010	Kurang	Dibuang / perlu direvisi
	0,011 - 0,050	Cukup	Baik
	0,051 - 1,000	Baik	Baik sekali
Reliabilitas	0,000 - 0,400	Rendah	Kurang baik
	0,401 - 0,700	Sedang	Cukup
	0,701 - 1,000	Tinggi	Baik

Dari hasil analisis butir soal di atas, dengan jumlah soal 45 butir dan sampel 24 orang, penafsiran terhadap beberapa soal tersebut antara lain seperti berikut.

Soal nomor 1, kunci jawabannya B, penafsirannya:

- Tingkat kesukaran butir soal (*Prop.Corrrect*) = 0,96 ; artinya butir soal ini 'sangat mudah' ($p > 0,901$) karena hampir 96 % dari peserta tes menjawab benar soal ini.
- Daya bedanya (*Biserial*) = 0,05 ; artinya daya beda yang 'sangat rendah' ($D > 0,119$).
- Option (alternatif jawaban) A dan B tidak berfungsi sebagai pengecoh (distraktor), karena seluruh peserta tes tidak ada satupun (0 %) yang memilih alterntif jawaban ini. Jika terjadi demikian, maka alternatif jawaban A dan B perlu diperbaiki karena tidak berfungsi sebagai pengecoh. Sesuai kriteria pengecoh yang 'baik' jika dipilih oleh 0,011 - 0,050 (1,1 % - 5 %).

Soal nomor 2, kunci jawabannya A, penafsirannya:

- Tingkat kesukaran butir soal (*Prop.Corrrect*) ini = 0,92 ; artinya butir soal ini 'mudah' ($p > 0,70$) karena lebih dari 92% peserta tes, menjawab benar soal ini.
- Daya bedanya (*Biserial*) = 0,46 ; artinya 'butir soal ini mempunyai daya beda yang 'tinggi' ($D > 0,40$)
- Option (alternatif jawaban) B dan C belum berfungsi sebagai pengecoh (distraktor), karena tidak ada yang memilih.

Soal nomor 6, kunci jawabannya B, penafsirannya:

- Analisis butir soal ini ada informasi '*Check The Key, b was specified, a works better*, artinya kunci jawaban a bekerja lebih baik daripada kunci jawaban yang telah ditentukan, yaitu b. Butir soal ini perlu dicek kembali, terbukti daya beda butir soal ini menunjukkan angka -0,33 (daya beda 'rendah' karena $D < 0,199$). Nilai daya beda yang demikian dapat diartikan, pada siswa pandai tidak dapat menjawab benar soal ini, tetapi siswa kurang pandai dapat menjawab benar (karena kebetulan saja).
- Tingkat kesukaran butir soal ini = 0,08 ; artinya butir soal ini 'sukar' ($p < 0,10$) karena hanya 8% dari peserta tes menjawab benar soal ini (pada b), dan yang menjawab pada a ada 88%. Soal nomor 6 ini soal ujian akhir sekolah (UAS) SD favorit di Bandar Lampung tahun 2004, yaitu:
FPB dan KPK bilangan 18 dan 48 adalah ...
a. 144 dan 6

- b. 12 dan 144
- c. 12 dan 288
- d. 288 dan 12

Jika dianalisis seluruh peserta, yaitu 219 orang, yang menjawab benar hanya 11% dan menjawab pada a ada 79%.

Bagaimana komentar Anda terhadap soal tersebut?

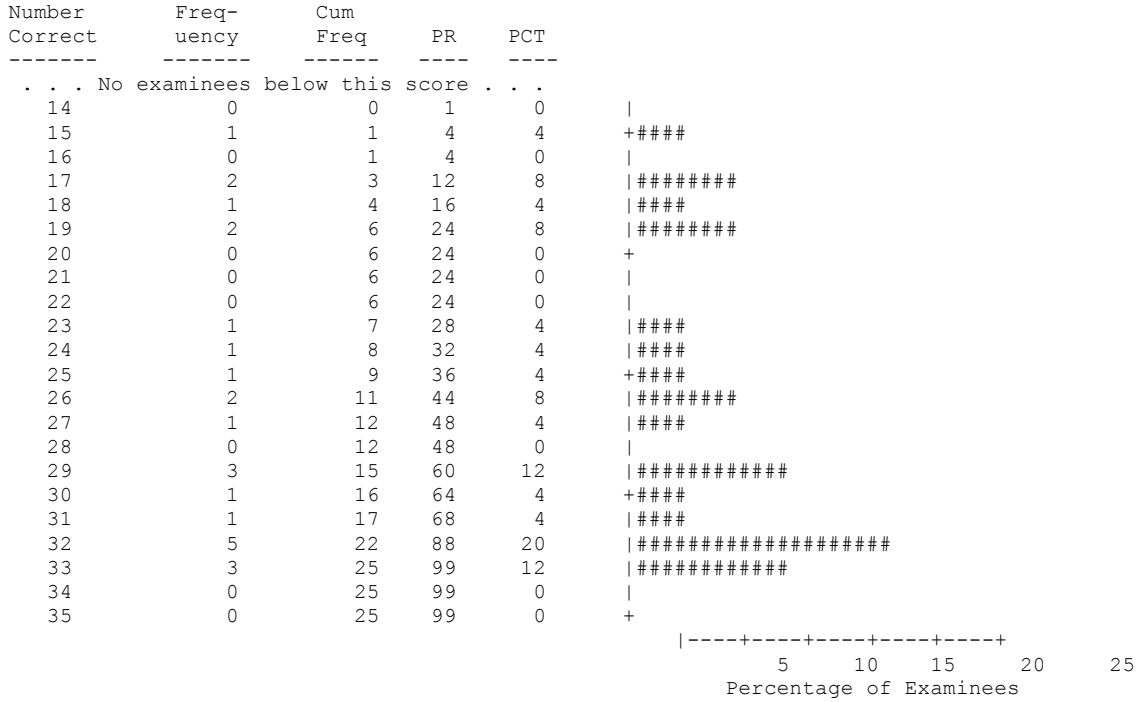
Scale Statistics, informasi penting yang dapat diperoleh antara lain:

1. N of Items (jumlah butir soal) = 35 → cek, apakah sudah benar
2. N of Examinees (jumlah peserta tes) = 25 → cek, apakah sudah benar ?
3. Skor setiap siswa, lihat pada file skor
4. Siswa banyak salah pada nomor berapa, lihat hasil analisis per butir soal
5. Pilihan jawaban yang tidak berfungsi sebagai pengecoh, lihat per butir soal
6. Mean (rerata skor) = 26,520 (keberhasilan 75,7% ; cukup baik)
7. Std. Dev. (Standard Deviasi) = 5,825
8. Minimum (skor terendah) = 15
9. Maximum (skor tertinggi) = 33
10. Alpha (nilai reliabilitas soal) = 0,874 → klasifikasi tinggi (soal ini baik)
11. Mean P (tingkat kesukaran soal) = 0,758 → klasifikasi mudah (perlu revisi)
12. Mean Bis. (daya beda soal) = 0,614 → klasifikasi tinggi (bagus sekali)

Grafik di bawah ini menunjukkan sebaran skor siswa, membacanya:

- Siswa memperoleh skor 15 ada 1 orang atau 4% (lihat kolom PR)
- Siswa memperoleh skor 33 ada 3 orang atau 12% (lihat kolom PR)
- Siswa memperoleh skor 18 dan dibawahnya (memperoleh nilai dibawah 6) ada 4 orang (lihat kolom Cum Freq)

Score Distribution Table (Tabel sebaran skor)



Daftar Pustaka

Dali S. Naga. 1992. *Pengantar teori sekora pada pengukuran pendidikan*. Jakarta: Besbats.

Hamid Hasan, S dan Asmawi Zainul. 1991. *Evaluasi hasil belajar*. Jakarta: Depdikbud.

Saifuddin Azwar. 1996. *Tes Prestasi*. Yogyakarta: Pustaka Pelajar.

Suharsimi Arikunto. 1993. *Dasar-dasar evaluasi pendidikan*. Jakarta: Bumi Aksara.

Sumadi Suryabrata. 1987. *Pengembangan tes hasil belajar*. Jakarta: Rajawali Pers.



ANALISIS BUTIR SOAL DENGAN KOMPUTER DAN
MENAFSIRKANNYA

Oleh
Drs. Ngadimun Hd, M.Pd.

Disampaikan dalam Sosialisasi KBK
bagi Guru SMP Kabupaten Tanggamus
di Pulau Panggung Tgl. 22-24 Juli 2004

BANDAR LAMPUNG
2004