

ANALISIS TES KINERJA DENGAN PARTIAL CREDIT MODEL (PCM)

Ni Made Sri Mertasari¹, I Made Candiasa²

^{1,2} Jurusan Matematika FMIPA UNDIKSA
Email: candiasa@undiksha.ac.id

ABSTRACT

Performance tests to evaluate character education in an integrated manner with subjects have been developed through research and are now trying to be trained for teachers in elementary schools through community service programs. Mastery of subjects is measured by the product produced, while character is measured by the process. The trainee teachers are invited to develop performance tests that can be used to measure subject mastery and character simultaneously. They were also asked to apply the test and correct the results using a rubric specially developed for this purpose. The rubric is developed following the steps of working on each item. Therefore, the rubric for assessing student answers varies for each item. The results of the test trials produced by the teachers were analyzed using the Partial Credit Model (PCM). PCM is a development of the Rasch model. The Rasch model applied to the dichotomous test items was developed into the PCM model applied to the polytomy test. Both the Rasch model and the PCM model use one parameter, namely item location or item difficulty level. The category score on the PCM indicates the number of steps to correctly complete the test item. By knowing the students' abilities, the probability of participants being able to answer the test items correctly can be estimated.

Keywords: Performance Test, Politymy, Partial Credit Model

ABSTRAK

Tes kinerja untuk mengevaluasi pendidikan karakter secara terpadu dengan mata pelajaran sudah dikembangkan melalui penelitian dan sekarang dicoba dilatihkan untuk guru-guru di Sekolah Dasar melalui program pengabdian kepada masyarakat. Penguasaan mata pelajaran diukur dari produk yang dihasilkan, sedangkan karakter diukur dari proses. Guru-guru peserta pelatihan diajak untuk mengembangkan tes kinerja yang dapat digunakan untuk mengukur penguasaan mata pelajaran dan karakter secara simultan. Mereka juga diminta menerapkan tes tersebut dan mengoreksi hasilnya dengan menggunakan rubrik yang khusus dikembangkan untuk tujuan tersebut. Rubrik dikembangkan mengikuti langkah-langkah pengerjaan masing-masing butir. Oleh karena itu, rubrik untuk menilai jawaban siswa bervariasi untuk masing-masing butir. Hasil uji coba tes yang dihasilkan para guru dianalisis menggunakan *Partial Credit Model (PCM)*. *PCM* merupakan pengembangan dari model *Rasch*. Model *Rasch* diterapkan untuk butir tes dikotomi dikembangkan menjadi model *PCM* yang diterapkan untuk tes politomi. Baik model *Rasch* maupun model *PCM* sama-sama menggunakan satu parameter, yakni lokasi butir atau tingkat kesukaran butir. Skor kategori pada *PCM* menunjukkan banyaknya langkah untuk menyelesaikan dengan benar butir tes tersebut. Dengan mengetahui kemampuan siswa, dapat diestimasi probabilitas peserta dapat menjawab butir tes dengan benar.

Kata-kata kunci: Tes Kinerja, Politomi, Partial Credit Model

PENDAHULUAN

Tes adalah suatu instrumen atau prosedur yang sistematis untuk mengukur suatu perilaku tertentu (Gronlund, 1985). Tes juga merupakan salah satu cara yang paling mudah yang bisa dilakukan untuk memotret kemajuan belajar peserta tes dalam ranah kognitif. Suatu tes

haruslah dapat mengukur hasil belajar yang sesuai dengan tujuan instruksional serta mempunyai reliabilitas dan validitas yang baik sehingga hasilnya dapat ditafsirkan dengan tepat. Tes dapat dibedakan menjadi dua macam, yakni tes subjektif dan tes objektif. Tes yang baik akan dapat mengestimasi dengan baik kemampuan peserta didik. Dalam dunia

pendidikan, kegiatan penilaian atau evaluasi hasil belajar peserta didik merupakan salah satu tugas penting yang harus dilakukan oleh pendidik. Evaluasi hasil belajar peserta didik dilakukan untuk mengetahui kemajuan peserta didik terhadap kurikulum yang telah diajarkan. Namun terkadang perangkat tes yang digunakan dalam mengukur kemampuan responden yang berbeda tidak mencerminkan prestasi belajar yang sebenarnya. Adapun solusi untuk mengatasi hal tersebut yaitu dengan menggunakan Teori Respon Butir.

Teori respon butir merupakan teori pengukuran modern yang digunakan dalam menganalisis butir. Teori ini mempunyai orientasi pada butir yang karakteristiknya tidak tergantung pada kelompok tertentu. Teori respon butir membebaskan ketergantungan antara butir tes dan peserta tes (konsep invariansi parameter), respon peserta tes pada satu butir tes tidak mempengaruhi butir tes lainnya (konsep independensi lokal), dan butir tes hanya mengukur satu dimensi ukur (konsep unidimensional). Dalam pelaksanaan tes kinerja, penskoran dilakukan perlangkah dan skor perbutir diperoleh peserta dengan menjumlah skor siswa tiap langkah, dan kemampuan diestimasi dengan skor mentah. Model penskoran seperti ini belum tentu tepat, karena tingkat kesulitan tiap langkah tidak diperhitungkan. Selain itu, peluang menjawab benar seorang siswa berdasarkan respon tertentu tidak dapat diprediksikan. Pendekatan alternatif yang dapat digunakan dengan pendekatan teori respon butir untuk penskoran politomi. Penskoran politomi adalah pemberian skor pada hasil tes yang terdiri dari dua nilai atau lebih, dimana penskorannya dilakukan langkah perlangkah dalam suatu butir soal, sehingga dalam proses analisisnya memperhitungkan tingkat kesulitan pada tiap langkah dalam menyelesaikan butir soal tersebut. Skor tertinggi tentu saja didapatkan ketika peserta tes mampu menyelesaikan dengan benar soal hingga langkah akhir.

Menurut Dali S.Naga (1992), Teori Respon Butir (*Butir Response Theory-IRT*) dikenal juga

sebagai Teori Ciri Laten (*Latent Trait Theory-LTT*) atau lengkungan karakteristik butir (*Butir Characteristic Function Curve-ICC*) atau Fungsi Karakteristik Butir (*Butir Characteristic Function-ICF*). Pada dasarnya, Teori Respon Butir ini ingin memperbaiki kelemahan yang terdapat pada teori tes klasik. Teori tes klasik merupakan teori pengukuran yang tertua di dunia, dimana teori tes klasik merupakan sebuah teori yang mudah dalam penerapannya serta model yang cukup berguna dalam mendeskripsikan bagaimana kesalahan dalam pengukuran dapat mempengaruhi skor amatan. Teori tes klasik dikenal juga dengan sebutan teori skor murni. Hal ini berkaitan dengan fokus kajian teori tes klasik yang ingin melihat skor murni dari skor tampak yang diperoleh. Pada Teori Respon Butir, peluang jawaban benar yang diberikan siswa, ciri atau parameter butir, dan ciri atau parameter peserta tes dihubungkan melalui suatu model formula yang harus ditaati baik oleh kelompok butir tes maupun kelompok peserta tes. Dalam hal ini, artinya butir yang sama terhadap peserta tes yang berbeda harus tunduk pada aturan rumus itu, atau peserta tes yang sama terhadap butir tes yang berbeda juga harus patuh terhadap rumus tersebut. Dalam proses seperti inilah dapat terjadi apa yang disebut dengan *invariansi* diantara butir tes dan peserta tes.

Menurut Hambleton, Swaminathan, & Rogers (1991), secara umum ciri-ciri Teori Respon Butir adalah sebagai berikut.

1. Karakteristik butir tidak tergantung pada peserta ujian.
2. Skor yang digambarkan peserta ujian tidak tergantung pada tes.
3. Merupakan model yang lebih menekankan pada tingkat butir dari pada tingkat tes.
4. Merupakan model yang tidak mensyaratkan secara ketat tes paralel untuk menaksir reliabilitas.
5. Merupakan model yang menguraikan sebuah ukuran keputusan untuk tiap skor kemampuan yakni ada hubungan fungsional antara peserta tes dengan tingkat kemampuan yang dimiliki.

Selain itu, Hambleton, Swaminathan, & Rogers (1991) mengemukakan bahwa asumsi-asumsi yang melandasi teori respon butir yaitu sebagai berikut.

1. Unidimensi, setiap butir hanya mengukur satu ciri peserta.
2. Independensi lokal, respon pada butir yang satu bebas dari pengaruh respon pada butir lain jika kemampuan yang mempengaruhi performansi dibuat konstan.
3. Fungsi karakteristik butir atau kurva karakteristik butir, merefleksikan hubungan yang sebenarnya antara kemampuan dan respon peserta terhadap butir tes.

Ada beberapa parameter yang umumnya dipertimbangkan dalam Teori Respon Butir, yaitu tingkat kesulitan butir (b), daya beda butir (a), peluang tebakan semu (c), parameter peserta (θ), dan respon peserta terhadap butir dinyatakan dalam bentuk probabilitas menjawab benar setiap langkah dalam butir ($P_i(\theta)$). PCM hanya menggunakan satu parameter, yaitu tingkat kesukaran butir (b). Menurut Arikunto (2010), indeks kesukaran (*difficulty index*) adalah bilangan yang menunjukkan sukar dan mudahnya suatu soal. Bilangan tersebut adalah bilangan real pada interval 0.00 sampai 1.00. Soal dengan indeks kesukaran 0.00 menunjukkan bahwa soal itu terlalu sukar, sebaliknya jika indeks kesukarannya 1.00 menunjukkan bahwa soalnya terlalu mudah. Perhitungan indeks tingkat kesukaran ini dilakukan untuk setiap nomor soal. Pada prinsipnya, skor rata-rata yang diperoleh peserta didik pada butir soal yang bersangkutan dinamakan tingkat kesukaran butir soal itu.

Fungsi tingkat kesukaran butir soal biasanya dikaitkan dengan tujuan tes. Misalnya, untuk keperluan ujian semester, digunakan butir soal yang memiliki tingkat kesukaran sedang, untuk keperluan seleksi digunakan butir soal yang memiliki tingkat kesukaran tinggi/sukar, dan untuk keperluan diagnosis biasanya digunakan butir soal yang memiliki tingkat kesukaran rendah/mudah. Dalam konteks ini, soal yang dinyatakan baik adalah soal yang

memiliki tingkat kesukaran yang sesuai dengan tujuan tes. Tingkat kesukaran soal dipandang dari kesanggupan atau kemampuan siswa dalam menjawabnya dan bukan dari sudut guru sebagai pembuat soal (Nana Sudjana, 2002). Formula indeks kesukaran untuk tes politomi adalah seperti berikut.

$$IK = \frac{\sum H + \sum L - (2 \cdot N \cdot Score_{\min})}{2 \cdot N (Score_{\max} - Score_{\min})}$$

Dalam formula di atas $\sum H$ dan $\sum L$ masing-masing menyatakan total skor kelompok atas dan total skor kelompok bawah, $Score_{\max}$ dan $Score_{\min}$ masing-masing menyatakan nilai maksimum dan nilai minimum, dan N menyatakan banyak data.

Menurut Ebel (1973) klasifikasi indeks kesukaran yang digunakan yaitu sebagai berikut.

Indeks Kesukaran	Kategori
0,00 – 0,30	Sukar
0,31 – 0,70	Sedang
0,71 – 1,00	Mudah

Persamaan Model Rasch dituliskan sebagai berikut.

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

Nilai peluang setiap peserta berhasil mengerjakan item i merupakan fungsi logistik perbedaan parameter kemampuan θ dengan tingkat kesukaran item b_i Persamaan di atas dapat ditulis kembali sebagai berikut.

$$P_i(\theta) = \frac{1}{1 + e^{-D(\theta - b_i)}} = \frac{P_{i1}}{P_{i0}(\theta) + P_{i1}(\theta)}$$

Persamaan RM untuk peserta n dan item i dengan skor x sebesar 0 dan 1 dengan kemampuan sebesar θ dan tingkat kesukaran item sebesar δ adalah sebagai berikut.

$$P_{nix} = \frac{1}{1 + \exp(\theta_n - \delta_{i1})} \text{ untuk } x = 0$$

$$P_{nix} = \frac{1}{1 + \exp(\theta_n - \delta_{i1})} \text{ untuk } x = 1$$

Model di atas merupakan model estimasi pada bentuk soal dengan skala penskoran dikotomi. Kemudian dikembangkan pada skala politomi yang memiliki pola skor x sebesar $x - 1$ dapat dihitung dengan persamaan sebagai berikut :

$$\frac{p_{ix}(\theta)}{P_{ix-1}(\theta) + P_{ix}(\theta)} = \frac{\exp(D(\theta - b_{ix}))}{1 + \exp(D(\theta - b_{ix}))} \text{ untuk } x = 0, 1, 2, 3, \dots, m_i$$

$P_{ix}(\theta)$ dan $P_{i(x-1)}(\theta)$ mengacu pada peluang seorang peserta dengan kemampuan θ meraih skor x dan $x - 1$. Peluang seorang peserta dengan kemampuan θ untuk memperoleh skor x pada item i dengan tingkat kesukaran item sebesar δ dapat dituliskan sebagai berikut.

$$P_{nix} = \frac{1}{1 + \exp(\theta_n - \delta_{ij})} \text{ untuk } x = 0$$

$$P_{nix} = \frac{\exp(\theta_n - \delta_{ij})}{1 + \exp(\theta_n - \delta_{ij})} \text{ untuk } x = 1, 2, 3, \dots, m_i$$

$P_{nix}(\theta)$ = peluang peserta ke- n dengan kemampuan θ dapat skor x pada butir i .

θ_n = tingkat kemampuan peserta

δ_{ij} = tingkat kesukaran tahap j pada butir i .

METODE

Tes kinerja untuk mengevaluasi pendidikan karakter secara terpadu dengan mata pelajaran sudah dikembangkan melalui penelitian dan sekarang dicoba dilatihkan untuk guru-guru di Sekolah Dasar melalui program pengabdian kepada masyarakat. Guru-guru diajak mengembangkan tes kinerja terpadu dengan mata pelajaran. Pada kesempatan ini, asesmen kinerja dibuat terpadu untuk mengukur karakter dan penguasaan materi matematika. Penguasaan mata pelajaran diukur dari produk yang dihasilkan, sedangkan karakter diukur dari proses. Masing-masing guru diminta untuk mengembangkan lima butir tes kinerja. Selanjutnya dilakukan diskusi untuk memilih lima tes terbaik untuk diterapkan.

Pada kesempatan berikutnya, guru-guru diminta menerapkan tes tersebut dan mengoreksi

hasilnya dengan menggunakan rubrik yang khusus dikembangkan untuk tujuan tersebut. Rubrik mengikuti prosedur penyelesaian tes kinerja. Masing-masing butir tes memiliki langkah pengerjaan yang berbeda. Oleh karena itu, masing-masing butir tes memiliki rubrik yang berbeda. Hasil uji coba tes yang dihasilkan para guru dianalisis menggunakan *Partial Credit Model (PCM)*. PCM merupakan pengembangan dari model *Rasch* butir dikotomi yang diterapkan pada butir politomi Widhiarso (2010). Skor kategori pada PCM menunjukkan banyaknya langkah untuk menyelesaikan dengan benar butir soal tersebut. Skor yang lebih tinggi menunjukkan kemampuan yang lebih besar daripada skor kategori yang lebih rendah. Jika i adalah butir politomi dengan kategori skor x sebesar $0, 1, 2, \dots, m_i$, maka probabilitas respon individu n dengan tingkat kemampuan θ memperoleh skor kategori j pada butir soal i , dengan persamaan, sebagai berikut.

$$\sum_{j=0}^0 (\theta - \delta_{ij}) = 0 \quad \text{dan}$$

$$\sum_{j=0}^h (\theta - \delta_{ij}) = \sum_{j=1}^h (\theta - \delta_{ij})$$

$$P_{nix}(\theta) = \frac{\exp \sum_{j=0}^x (\theta - \delta_{ij})}{\sum_h \exp \sum_{j=0}^h (\theta - \delta_{ij})} \quad \text{untuk } x = 0, 1, 2, \dots, m_i$$

dimana:

$P_{nix}(\theta)$ = Peluang peserta kemampuan θ memperoleh skor x pada butir i

x = Skor peserta

j = Kategori/tahap dalam butir soal

i = Butir soal

n = Peserta

θ_n = Tingkat kemampuan peserta ke n

δ_{ij} = Tingkat kesukaran tahap (j) pada butir (i)

Langkah-langkah analisis butir soal dengan menggunakan PCM adalah sebagai berikut.

1. Input skor yang diperoleh responden dengan penskoran politomi
2. Urutkan data penskoran dari skor yang tertinggi sampai skor yang terendah berdasarkan total skor yang diperoleh responden
3. Tentukan kelompok atas dan kelompok bawah, dimana jika jumlah responden kurang dari 30 orang maka untuk mencari kelompok atas dan kelompok bawah tinggal mencari 50% dari jumlah peserta. Jika jumlah responden lebih dari 30 orang maka diambil 27% untuk kelompok atas dan 27% untuk kelompok bawah.
4. Ubah skor mentah tersebut menjadi skor politomi dengan 5 kategori untuk menunjukkan banyaknya langkah dalam pengujian PCM pada tiap butir soal.
5. Hitung tingkat kesukaran butir soal.
6. Hitung peluang probabilitas dengan PCM.

HASIL DAN PEMBAHASAN

Masing-masing guru peserta diminta untuk menyusun lima butir tes kinerja. Selanjutnya dilakukan diskusi untuk memilih lima butir tes terbaik. Lima butir tes kinerja yang dihasilkan peserta dalam bentuk tes kinerja kemudian diujicoba di kelas. Hasil uji coba kemudian dikoreksi menggunakan rubrik yang dikembangkan khusus untuk tujuan tersebut. Masing-masing butir memiliki langkah pengerjaan yang berbeda. Oleh karena itu, rubrik untuk masing-masing butir juga berbeda. Skor tertinggi ideal yang dapat diraih peserta adalah 100 dan skor terendah ideal yang mungkin diperoleh peserta adalah 0. Tes tersebut diujicobakan terhadap 20 siswa. Hasil yang diperoleh adalah seperti tercantum pada Tabel-1.

Agar dapat menerapkan PCM, skor mentah pada Tabel-1 dikonversi dalam bentuk skor politomi. Sesuai dengan rubrik penskoran yang ada, skor untuk butir soal ke-1 akan dikonversi dengan 9 kategori yaitu 1, 2, 3, 4, 5, 6, 7, 8, dan 9. Skor untuk butir soal ke-2 akan dikonversi dengan 8 kategori yaitu 1, 2, 3, 4, 5, 6, 7, dan 8.

Skor untuk butir soal ke-3 akan dikonversi dengan 6 kategori yaitu 1, 2, 3, 4, 5, dan 6. Skor untuk butir soal ke-4 akan dikonversi dengan 5 kategori yaitu 1, 2, 3, 4, dan 5. Skor untuk butir soal ke-5 akan dikonversi dengan 7 kategori yaitu 1, 2, 3, 4, 5, 6, dan 7. Langkah selanjutnya yaitu menghitung tingkat kesukaran dari masing-masing butir soal dengan menggunakan rumus tingkat kesukaran untuk soal politomi. Peluang siswa dalam menjawab butir soal dengan benar menggunakan formula PCM.

Tabel-1: Hasil Uji Coba

Resp.	Nomor Butir					Total
	1	2	3	4	5	
001	100	100	100	100	100	500
002	100	100	100	85	100	485
003	100	100	67	97	100	464
004	65	100	67	95	95	422
005	100	60	100	85	37	382
006	100	100	37	40	100	377
007	50	65	67	55	100	337
008	20	100	67	70	70	327
009	50	35	60	85	50	280
010	50	100	100	15	5	270
011	100	5	27	5	100	237
012	30	45	40	15	100	230
013	70	100	15	37	5	227
014	25	60	100	23	15	223
015	15	100	35	20	30	200
016	50	20	5	5	100	180
017	100	52	0	20	5	177
018	65	20	30	45	10	170
019	50	15	40	0	5	110
020	50	0	35	0	5	90

Berikut diuraikan perhitungan untuk butir nomor-1. Mula-mula dihitung tingkat kesukaran butir.

$$\delta_{ij} = \frac{\sum H + \sum L - (2 \cdot N \cdot Score_{min})}{2 \cdot N (Score_{max} - Score_{min})}$$

$$\delta_{ij} = \frac{735 + 555 - (2 \cdot 20 \cdot 15)}{2 \cdot 20 (100 - 15)} = 0,20$$

Tingkat kemampuan responden diasumsikan berkemampuan standar yaitu 0,5. Dengan demikian diperoleh peluang responden seperti berikut.

$$P_{nix}(\theta) = \frac{\exp \sum_{j=0}^x (\theta - \delta_{ij})}{\sum_h^{m_i} \exp \sum_{j=0}^h (\theta - \delta_{ij})}$$

$$P_{20,1,9}(0,5) = \frac{\exp \sum_{j=1}^9 (0,5 - 0,20_j)}{\sum_h^9 \exp \sum_{j=1}^h (0,5 - 0,20_j)}$$

- $\exp \sum_{j=1}^1 (0,5 - (0,20_j)) = \sum_1^9 1,35$
- $\exp \sum_{j=1}^2 (0,5 - (0,20_j)) = \sum_2^9 1,5$
- $\exp \sum_{j=1}^3 (0,5 - (0,20_j)) = \sum_3^9 1,35$
- $\exp \sum_{j=1}^4 (0,5 - (0,20_j)) = \sum_4^9 1$
- $\exp \sum_{j=1}^5 (0,5 - (0,20_j)) = \sum_5^9 0,61$
- $\exp \sum_{j=1}^6 (0,5 - (0,20_j)) = \sum_6^9 0,3$
- $\exp \sum_{j=1}^7 (0,5 - (0,20_j)) = \sum_7^9 0,12$
- $\exp \sum_{j=1}^8 (0,5 - (0,20_j)) = \sum_8^9 0,04$
- $\exp \sum_{j=1}^9 (0,5 - (0,20_j)) = \sum_9^9 0,01$

$$P_{20,1,9}(0,5) = \frac{\exp \sum_{j=1}^9 (0,5 - 0,20_j)}{\sum_h^9 \exp \sum_{j=1}^h (0,5 - 0,20_j)}$$

$$= \frac{\exp (1,1)}{6,28} 0,478.$$

Berdasarkan hasil di atas, didapat bahwa peserta tes dengan tingkat kemampuan (θ) sama dengan 0,5 memiliki peluang untuk menjawab butir nomor-1 dengan benar sebesar 0,478. Dengan prosedur yang sama, diperoleh peluang responden dengan kemampuan (θ)=0,5 untuk menjawab setiap butir dengan benar seperti tercantum pada Tabel-2 berikut.

Tabel-2: Hasil Analisis

No. Butir	Indeks Kesukaran (b)	Peluang Menjawab Benar $P(\theta)$
1	0,20	0,478
2	0,25	0,482
3	0,32	0,521
4	0,14	0,233
5	0,34	0,532

Tujuan dari pengujian butir tes di atas adalah untuk mengetahui apakah butir soal memenuhi syarat untuk diterapkan. Berdasarkan teori PCM ini, butir tes yang memenuhi syarat untuk diterapkan adalah butir tes yang layak. Butir tes dianggap layak apabila tingkat kemampuan peserta tes diambil 0,5, maka peluang $P_{nix}(\theta)$ peserta tes untuk memperoleh skor tertinggi terletak antara 0,45 dan 0,55. Hal tersebut didapatkan berdasarkan toleransi peluang dan kriteria pengujiannya yang ditentukan berdasarkan nilai alpha yang digunakan yaitu 0,5. Berdasarkan analisis dan perhitungan yang

lakukan diperoleh ada satu butir tes yang tidak memenuhi syarat yaitu butir nomor 4.

SIMPULAN

PCM merupakan perluasan dari Rasch Model, sehingga diasumsikan bahwa semua butir soal mempunyai daya beda yang sama. Pada PCM diasumsikan bahwa parameter tingkat kesulitan butir merupakan satu-satunya karakteristik butir yang mempengaruhi kinerja peserta didik. Peluang peserta didik berhasil mengerjakan butir soal tergantung pada kemampuan dan tingkat kesulitan butir tes yang dikerjakannya. PCM tidak mensyaratkan langkah penyelesaian butir tes harus berurutan dan tidak harus mempunyai kesulitan yang sama. Pengujian dengan PCM menggunakan skor kategori yang bertujuan untuk menentukan banyaknya langkah dalam menyelesaikan dengan benar butir soal tertentu sehingga kemampuan setiap peserta tes dapat diestimasi dengan menghitung probabilitas setiap peserta tes dalam menyelesaikan sebuah soal tes.

Ada lima butir tes kinerja yang dihasilkan guru-guru peserta pelatihan yang dianalisis menggunakan PCM. Formula untuk menentukan probabilitas adalah sebagai berikut.

$$P_{nix}(\theta) = \frac{\exp \sum_{j=0}^x (\theta - \delta_{ij})}{\sum_h^{m_i} \exp \sum_{j=0}^h (\theta - \delta_{ij})} \quad \text{untuk } x = 0, 1, 2, \dots, m_i$$

$P_{nix}(\theta)$ = Peluang peserta memperoleh skor x pada butir i (kemampuan = θ).

x = Skor peserta

j = Kategori/tahap dalam butir soal

i = Butir soal

n = Peserta

θ_n = Tingkat kemampuan peserta ke n

δ_{ij} = Tingkat kesukaran tahap j pada butir i

Tingkat kemampuan peserta tes diambil 0,5, sehingga toleransi peluang ditentukan berdasarkan nilai alpha yang digunakan yaitu 0,5. Dengan demikian butir tes dianggap layak apabila peluang $P_{nix}(\theta)$ peserta tes untuk memperoleh skor tertinggi terletak antara 0,45 dan 0,55. Jadi butir tes yang memenuhi syarat kelayakan adalah butir nomor-1 dengan peluang 0,478; butir nomor-2 dengan peluang 0,462; butir nomor-3 dengan peluang 0,521, dan butir nomor-5 dengan peluang 0,532. Butir nomor-4

dengan peluang 0,233 tidak memenuhi syarat kelayakan.

DAFTAR PUSTAKA

- Arikunto, Suharsimi. 2010. *Dasar – Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara
- Candiasa, I Made. 2010a. *Pengujian Instrumen Penelitian Disertai Aplikasi ITEMAN dan BIGSTEPS*. Singaraja: Unit Penerbitan Universitas Pendidikan Ganesha.
- Ebel, Robert L., *Essential of Educational Measurement*, Englewood Cliff, NJ: Prentice Hall, Inc., 1972
- Gronlund, Norman E. 1985. *Measurement and Evaluation in Teaching*. MacMillan Publishing Company. New York.
- Hambleton R.K., Swaminathan H., and Rogers H.J., 1991. *Fundamental of Butir Response Theory*. Sage Publication Inc., Newbury Park, CA.
- Naga, Dali S. “Estimasi Kemampuan Peserta Ujian Melalui Karakteristik Butir Model Logistik.” *Buletin Ilmiah Tarumanagara*. Th. 7, No. 24
- Sudijono, Anas. 2012. *Pengantar Evaluasi Pendidikan*. Jakarta: Raja Grafindo Persada
- Sudjana, Nana. 2002. *Dasar-dasar Proses Belajar Mengajar*. Bandung: Sinar Baru Algensindo
- Widhiarso W., 2010. *Model Politomi dalam Teori Respon Butir*. Fakultas Psikologi UGM, Yogyakarta.