

## MENGESTIMASI PARAMETER BUTIR PERANGKAT TES DENGAN PENDEKATAN REGRESI LOGISTIK MULTIVARIAT

**Heri Retnawati**

*Pendidikan Matematika FMIPA UNY  
(retnawati\_heriuny@yahoo.co.id)*

### **Abstrak**

Estimasi parameter butir perangkat tes biasanya dilakukan dengan pendekatan teori tes klasik atau teori respons butir (*Item Response Theory, IRT*) unidimensi. Untuk analisis ini, ada asumsi yang harus dipenuhi, salah satunya asumsi unidimensi, yakni suatu perangkat tes dianggap hanya mengukur satu kemampuan yang paling dominan saja. Pada kenyataannya di lapangan, asumsi ini sulit dipenuhi, karena sebagian besar tes kemampuan/psikologi bersifat multidimensi. Teori respons butir multidimensi mengatasi kelemahan ini, yang mengestimasi parameter butir melalui persamaan peluang menjawab benar dengan regresi logistik multivariat.

Kata kunci : *IRT*-unidimensi, *IRT*-multidimensi

### **PENDAHULUAN**

Untuk mengetahui kemampuan seseorang di bidang tertentu, tes merupakan salah satu instrumen yang sering digunakan. Tes dapat digunakan setelah melalui proses pengembangan. Pengembang tes dapat melakukan perancangan, perakitan dan analisis butir tes dengan menggunakan berbagai pendekatan, baik menggunakan teori tes klasik maupun teori respons butir. Pengembangan tes dengan pendekatan teori respons butir memiliki keunggulan. Model teori respons butir menghasilkan parameter butir yang independen dari peserta tes dan parameter peserta tes yang bersifat independen dari sekumpulan butir yang diujikan (Xitao Fan, 1998). Sifat invarian dari model teori respons butir membuat teori respons butir secara teoretis bisa digunakan untuk menyelesaikan permasalahan yang tidak dapat diselesaikan oleh teori tes klasik. Pada teori respons butir, ada asumsi yang harus dipenuhi, yakni independensi lokal dan unidimensi (Hambleton, Swaminathan, dan Rogers, 1991).

Independensi lokal terjadi jika kemampuan yang mempengaruhi kinerja suatu tes adalah konstan, artinya respons peserta tes dalam menjawab suatu butir tes independen secara statistik terhadap respons peserta tes tersebut dalam menjawab butir lainnya. Dengan kata lain, asumsi independensi lokal menyatakan bahwa tidak ada korelasi antara respons peserta tes pada butir soal yang berbeda. Ini berarti bahwa kemampuan yang dinyatakan dalam model merupakan satu-satunya faktor yang mempengaruhi respons peserta tes terhadap butir soal. Unidimensi, artinya tes hanya mengukur satu kemampuan tertentu. Asumsi unidimensi dapat ditunjukkan jika tes mengandung hanya satu komponen dominan yang mengukur prestasi suatu subjek. Pada praktiknya, asumsi unidimensi tidak dapat dipenuhi secara ketat karena adanya faktor-faktor kognitif, kepribadian, dan administratif dalam tes, seperti kecemasan, motivasi, dan tendensi untuk menebak. Pada kenyataannya di lapangan, asumsi unidimensi sulit terpenuhi. Hal ini sesuai dengan pendapat bahwa kebanyakan tes pendidikan dan psikologi pada beberapa tingkat bersifat multidimensi (Bolt dan Lall, 2003; Ackerman, dkk., 2003).

Pada data yang realitasnya multidimensi, analisis dengan model unidimensi akan menyesatkan. Terkait dengan hal ini, pada data yang bersifat multidimensi, diperlukan analisis dengan pendekatan multidimensi untuk mengestimasi parameter peserta dan parameter butir. Pada tulisan ini dibahas tentang parameter peserta, parameter butir tes, dan estimasinya dengan pendekatan teori respons butir multidimensi.

## PEMBAHASAN

Pada teori respons butir unidimensi, hubungan antara tiga parameter butir yaitu indeks kesukaran butir soal, indeks daya beda butir, dan indeks tebakan semu (*pseudoguessing*) dan satu kemampuan yang dinyatakan dalam persamaan peluang menjawab benar. Secara matematis, model logistik tiga parameter dapat dinyatakan sebagai berikut (Hambleton, dan Swaminathan, 1985 : 49; Hambleton, Swaminathan, dan Rogers, 1991: 17).

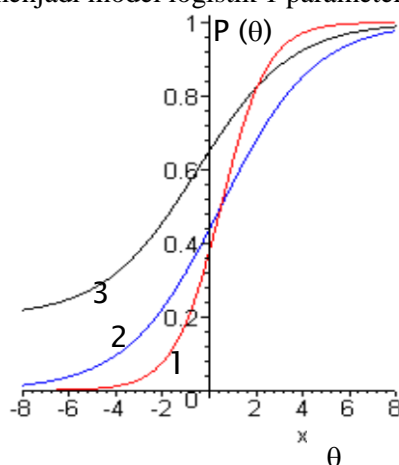
$$P_i(\theta) = c_i + (1-c_i) \frac{e^{Da_i(\theta-b_i)}}{1 + e^{Da_i(\theta-b_i)}} \dots\dots\dots (1)$$

Keterangan :

- $\theta$  : tingkat kemampuan peserta tes
- $P_i(\theta)$  : probabilitas peserta tes yang memiliki kemampuan  $\theta$  dapat menjawab butir  $i$  dengan benar
- $a_i$  : indeks daya beda dari butir ke- $i$
- $b_i$  : indeks kesukaran butir ke- $i$
- $c_i$  : indeks tebakan semu butir ke- $i$
- $e$  : bilangan natural yang nilainya mendekati 2,718
- $n$  : banyaknya item dalam tes
- $D$  : faktor penskalaan yang harganya 1,7.

Parameter  $b_i$  merupakan suatu titik pada skala kemampuan dalam kurva karakteristik butir ketika peluang menjawab benar peserta tes sebesar 50%. Parameter  $a_i$  merupakan indeks daya pembeda yang dimiliki butir ke- $i$ . Pada kurva karakteristik,  $a_i$  proporsional terhadap koefisien arah garis singgung (*slope*) pada titik  $\theta = b$ . Butir soal yang memiliki daya pembeda yang besar mempunyai kurva yang sangat menanjak, sedangkan butir soal yang mempunyai daya pembeda kecil mempunyai kurva yang sangat landai. Parameter ini menggambarkan probabilitas peserta dengan kemampuan rendah menjawab dengan benar pada suatu butir. Dengan adanya indeks tebakan semu pada model logistik tiga parameter, memungkinkan subjek yang memiliki kemampuan rendah mempunyai peluang untuk menjawab butir soal dengan benar. Nilai kemampuan peserta ( $\theta$ ) biasanya terletak di antara  $-3$  dan  $+3$ , sesuai dengan daerah asal distribusi normal. Gambar 1 menyajikan kurva karakteristik butir 1 ( $a=1, b=0,5, c=0$ ), butir 2( $a=0,5, b=0,5, c=0$ ) dan butir 3 ( $a=0,5, b=0,5, c=0,2$ ).

Model logistik 2 parameter dan model 1 parameter merupakan kasus dari model logistik 3 parameter. Ketika indeks *pseudo-guessing* sama dengan 0 ( $c=0$ ), model tersebut menjadi model 2 parameter. Demikian pula pada model 2 parameter, ketika indeks daya pembeda butir bernilai 1, maka model ini menjadi model logistik 1 parameter, atau lebih dikenal dengan nama model Rasch.



Gambar 1. Kurva karakteristik butir model 3P, dengan butir 1 ( $a=1, b=0,5, c=0$ ), butir 2( $a=0,5, b=0,5, c=0$ ) dan butir 3 ( $a=0,5, b=0,5, c=0,2$ )

Seperti halnya teori respons butir unidimensi, pada model teori respons butir multidimensi data dapat berupa butir skor dikotomi atau politomi. Matriks data disusun sedemikian rupa, dengan  $x_{ij}$  menyatakan elemennya pada baris ke- $i$  dan kolom ke- $j$ . Butir dinyatakan dalam  $i$  ( $i=1, \dots, n$ ) dan peserta dinyatakan dalam  $j$  ( $j=1, \dots, N$ ).

Dalam menyusun matriks data, ada asumsi yang harus diperhatikan (Reckase, 1997), yakni :

- a. Semakin tinggi kemampuan peserta tes, probabilitas menjawab benar peserta tes terhadap butir tes tidak menurun (asumsi kemonotonan).
- b. Fungsi probabilitas menjawab benar bersifat *smooth* (turunan fungsinya terdefiniskan).
- c. Probabilitas kombinasi respons dapat ditentukan dengan hasil probabilitas respons individual ketika probabilitas dihitung kondisional pada titik dalam ruang yang didefinisikan oleh konstruk hipotetik (asumsi independensi lokal).

Berdasarkan hal ini, asumsi yang digunakan untuk menyusun matriks data yakni asumsi kemonotonan, memiliki turunan fungsi, dan independensi lokal.

Pada teori respons butir multidimensi (*multidimensional item response theory, MIRT*) dikenal dua model, yakni *compensatory* dan *noncompensatory*. Menurut Ansley dan Forsyth (Spray, dkk., 1990), model *compensatory* membolehkan kemampuan tinggi pada salah satu dimensi memperoleh kompensasi pada kemampuan rendah pada dimensi lain dalam kaitannya dengan probabilitas menjawab benar. Sebaliknya, pada model *noncompensatory* tidak membolehkan kemampuan tinggi pada salah satu memperoleh kompensasi pada kemampuan rendah pada dimensi lainnya. Untuk model *compensatory* pada kasus butir dua dimensi, seorang peserta tes dengan kemampuan sangat rendah pada satu dimensi dan kemampuannya sangat tinggi pada dimensi lain dapat menjawab butir tes dengan benar.

Ada dua tipe model *compensatory*, yakni model MIRT logistik (Reckase, 1997) dan model *ogive* normal dari Samejima dengan menyatakan kombinasi linear dari kemampuan multidimensi dalam pangkat pada rumus probabilitas menjawab benar. Dalam model linear ini, rendahnya satu atau lebih kemampuan dapat dikompensasikan pada dimensi lainnya. Karena kompensasi merupakan karakteristik kombinasi linear, maka model ini diberi nama dengan model MIRT linear (Spray, dkk., 1990; Bolt dan Lall, 2003) yang merupakan regresi logistik multivariat. Model MIRT logistik linear dapat ditulis sebagai :

$$P_i(\theta_j) = c_i + (1-c_i) \frac{e^{\sum_{m=1}^k f_{ijm} + d_i}}{(1 + e^{\sum_{m=1}^k f_{ijm} + d_i})} \dots\dots\dots(2)$$

Dengan  $f_{ijm} = \mathbf{a}_{jm} \cdot \theta_{im}$ ,  $c_i$  merupakan parameter *pseudo-guessing* butir ke- $i$ ,  $\mathbf{a}_{jm}$  parameter diskriminasi untuk butir ke- $i$  pada dimensi ke- $m$ ,  $d_i$  parameter tingkat kesulitan butir ke- $i$ , dan  $\theta_{jm}$  merupakan elemen ke- $m$  dari vektor kemampuan orang ke  $j$  ( $\theta_j$ ). Senada dengan itu, Kirisci, Hsi dan Yu (2001) menuliskan persamaan (2) sebagai

$$P_i(\theta_j) = c_i + \frac{1 - c_i}{(1 + e^{-1.7(a_j \theta_j - b_i)})} \dots\dots\dots(3)$$

yang memuat skala penyekalaan  $D=1,7$ .

Di lain pihak, model MIRT *noncompensatory* dideskripsikan sebagai probabilitas dari respons yang menguntungkan pada hasil kali dari fungsi kemampuan sebanyak  $k$  dimensi dan karakteristik butir. Model MIRT logistik tipe *noncompensatory* dapat ditulis sebagai :

$$P_i(\theta_j) = c_i + (1-c_i) \prod_{m=1}^k \frac{e^{f_{ijm}}}{(1 + e^{f_{ijm}})} \dots\dots\dots(4)$$

Dengan  $f_{ijm} = [a_{im} (\theta_{jm} - b_{im})]$  dengan  $b_{im}$  merupakan parameter tingkat kesulitan butir ke-i pada dimensi ke-m. Terkait dengan bentuknya yang merupakan hasil perkalian, model ini sering pula dinamai dengan model multiplikatif.

Mengingat pada tulisan ini lebih difokuskan pada MIRT model *compensatory*, maka hanya model linear ini saja yang akan dibahas. Seperti halnya pada teori respons butir model 3 parameter, parameter-parameter model ini meliputi parameter peserta tes, daya pembeda, tingkat kesulitan dan tebakan semu.

Parameter peserta tes pada model ini dinyatakan dengan elemen-elemen dari vektor  $\theta_j$ . Banyaknya elemen dari vektor ini masih merupakan hal yang sering diperdebatkan (Reckase, 1997). Berdasarkan pengalaman Reckase dan Hirsch (Reckase, 1997), banyaknya dimensi kemampuan sering *underestimate* dan *overestimate* dan hal ini akan merugikan. Banyaknya dimensi yang digunakan pada model tergantung interaksi butir dengan peserta tes yang perlu disesuaikan dengan tujuan analisis.

Diskriminasi butir pada teori respons butir multidimensi merupakan parameter untuk model yang dinyatakan dengan vektor  $\mathbf{a}$  yang fungsinya mirip dengan parameter  $a$  pada teori respons butir unidimensi. Unsur-unsur vektor terkait dengan kemiringan dari permukaan respons pada arah yang bersesuaian dengan sumbu- $\theta$ . Kemiringan ini mengindikasikan sensitivitas butir terhadap kemampuan sepanjang sumbu- $\theta$ . Jika parameter ini mengukur bukan hanya satu dimensi saja, maka diskriminasi butir dapat dinyatakan dengan kombinasi dimensi-dimensi, yang dinyatakan dengan

$$MDISC_i = \sqrt{\sum_{m=1}^k a_{im}^2} \dots\dots\dots(5)$$

$MDISC_i$  merupakan diskriminasi dari butir  $i$  untuk kombinasi terbaik dari kemampuan,  $k$  banyaknya dimensi pada ruang- $\theta$ , dan  $a_{im}$  merupakan elemen dari vektor  $\mathbf{a}$  untuk butir ke- $i$ .

Tingkat kesulitan butir merupakan parameter  $d_i$  pada model. Parameter ini tidak dapat diinterpretasikan dengan cara yang sama dengan parameter- $b$  pada teori respons butir unidimensi. Misalkan  $a$  merupakan parameter daya pembeda butir pada model unidimensi, maka  $-ab = d_i$ . Nilai yang ekuivalen dengan tingkat kesulitan pada model unidimensi dinyatakan dengan

$$MDIFF_i = \frac{-d_i}{MDISC} \dots\dots\dots(6)$$

$MDIFF_i$  menyatakan jarak dari titik asal ruang- $\theta$  terhadap titik kemiringan paling tinggi pada arah dari titik asal. Menurut Ackerman, dkk. (2003), tanda dari jarak ini mengindikasikan kesulitan relatif butir. Sebagai contoh, pada tes yang memuat dua dimensi, butir dengan  $MDIFF_i$  negatif, relatif mudah dan berada di kuadran III; dan relatif sulit jika terletak di kuadran I.  $MDIFF_i$  analog dengan parameter  $b$  pada teori respons butir unidimensi. Lokasi (parameter lokasi) bersesuaian dengan arah sudut butir dari tiap butir relatif terhadap sumbu  $\theta_i$  positif. Arah kemiringan yang paling besar dari titik pusat koordinat, menurut Reckase (1997) dan Ackerman, dkk. (2003) dinyatakan dengan

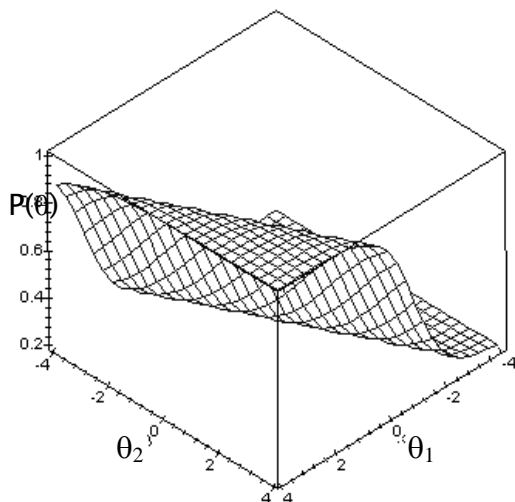
$$\alpha_i = \arccos \frac{a_{im}}{MDSC_i} \dots\dots\dots(7)$$

dengan  $a_{im}$  merupakan sudut antara garis dari titik pusat koordinat ke titik yang memiliki kemiringan terbesar dengan sumbu ke- $m$  untuk butir ke- $i$ .

Asimtot yang lebih rendah (*lower asymptote*) merupakan nilai yang menyatakan probabilitas menjawab benar peserta tes ketika kemampuan yang dimilikinya sangat rendah pada keseluruhan dimensi. Parameter ini sama artinya dengan parameter  $c$  pada teori respons butir unidimensi.

Jika dibandingkan dengan model logistik pada teori respons butir unidimensi, perbedaan ini akan sangat mencolok dengan mencermati kurva karakteristik butir pada model logistik multidimensi. Sebagai ilustrasi, pada Gambar 4 disajikan kurva karakteristik butir yang memuat dua dimensi, dengan parameter  $a_1=1, a_2=1, d = 1, c=0,2$ . Kurva karakteristik butir pada model ini nampak sebagai permukaan, sehingga sering disebut pula dengan permukaan respons butir (*Item Response Surface, IRS*) (Bolt dan Lall, 2003) atau permukaan karakteristik butir (*Item*

Characteristic Surface, ICS) (Ackerman, dkk., 2003). Permukaan respons butir ini akan sangat sulit digambarkan jika dimensi kemampuan yang diukur suatu butir tes lebih dari dua.



Gambar 2. Permukaan karakteristik butir yang mengukur dua dimensi, dengan parameter  $a_1=1$ ,  $a_2=1$ ,  $d = 1$ ,  $c=0,2$

Fungsi informasi butir pada teori respons butir multidimensi dinyatakan dengan :

$$I_{i\alpha}(\theta) = \frac{[\nabla_{\alpha} P_i(\theta)]^2}{P_i(\theta)[1 - P_i(\theta)]} \dots\dots\dots(8)$$

Dengan  $I_{i\alpha}(\theta)$  merupakan informasi yang disajikan oleh butir-i pada arah  $\alpha$  dalam ruang dan  $\nabla_{\alpha}$  merupakan operator definitif untuk turunan dengan arah  $\alpha$ . DeBryant (tth) menyajikan fungsi informasi butir berarah ini secara lebih detail. Fungsi informasi berarah disajikan sebagai :

$$I_i(\theta_j) = D^2 (\mathbf{a}_i' \mathbf{u}_i)^2 Q_i(\theta_j) \{ P_i(\theta_j) [1 + \text{Exp}(-L)]^2 \}^{-1} \dots\dots\dots(9)$$

dengan  $L = D (\mathbf{a}_i' \theta_j + d_i)$ . Skala kemampuan yang memaksimumkan nilai fungsi informasi yakni

$$\theta_{\max} = \mathbf{u}_i [\ln \{ .5 [1 + (8c_i + 1)^{1/2}] \} (D' \mathbf{a}_i)^{-1} - d_i ( \mathbf{a}_i )^{-1} ] \dots\dots\dots(10)$$

Atau jika dinyatakan dengan tingkat kesulitan (MDIFF) dan indeks daya pembeda butir (MDISC) menjadi

$$\theta_{\max} = \mathbf{u}_i [\ln \{ .5 [1 + (8c_i + 1)^{1/2}] \} (D.MDISC_i)^{-1} + MDIFF_i ] \dots\dots\dots(11)$$

Sebagai akibatnya, kemampuan yang memaksimumkan IIF pada dimensi ke-m yakni

$$\theta_{\max m} = [\ln \{ .5 [1 + (8c_i + 1)^{1/2}] \} (D.MDISC_i)^{-1} + MDIFF_i ] \cos \alpha_{mi} \dots\dots\dots(12)$$

Pada teori respons butir multidimensi, parameter-parameter pada model dapat diestimasi dengan menggunakan berbagai prosedur, misalnya likelihood gabungan maksimum (*joint maximum likelihood procedures*) (Reckase, 1997). Prosedur estimasi ini bertujuan untuk menemukan himpunan parameter butir dan peserta tes yang akan memaksimumkan likelihood dari respons butir yang teramati. Bentuk persamaan likelihoodnya diberikan oleh :

$$Li = \prod_{j=1}^N \prod_{i=1}^n P(u_{ij} | a_i, d_i, c_i, \theta_j) \dots\dots\dots(13)$$

dengan  $u_{ij}$  merupakan respons butir-i oleh orang j, baik 0 ataupun 1. Terkait dengan perhitungan secara matematis, biasanya program komputer yang ada meminimumkan negatif logaritma dari Li atau  $F = - \ln(Li)$ .

Pengujian kegunaan model multidimensi yang diusulkan dilakukan menganalisis kecocokan model (*goodness of fit*). Salah satu prosedur yang dapat digunakan adalah cara yang disarankan

Reckase (1997). Cara ini ditempuh dengan menguji unsur-unsur matriks korelasi residual antarbutir. Korelasi residual antara butir  $i_1$  dan  $i_2$ , dinyatakan dengan  $r_{i_1 i_2}$  yang dihitung dengan :

$$r_{i_1 i_2} = \frac{1}{n} \sum_{a=1}^N \frac{(u_{i_1 a} - P_{i_1 a})(u_{i_2 a} - P_{i_2 a})}{\sqrt{P_{i_1 a} Q_{i_1 a}} \sqrt{P_{i_2 a} Q_{i_2 a}}} \dots\dots\dots(14)$$

dengan  $P_{ia}$  merupakan probabilitas kondisional model tersimplifikasi dari jawaban benar butir  $i$  untuk peserta tes yang terpilih secara acak. Kecocokan model yang baik menghasilkan residu estimasi korelasi hasil observasi antarbutir mendekati 0 ketika  $N$  besar.

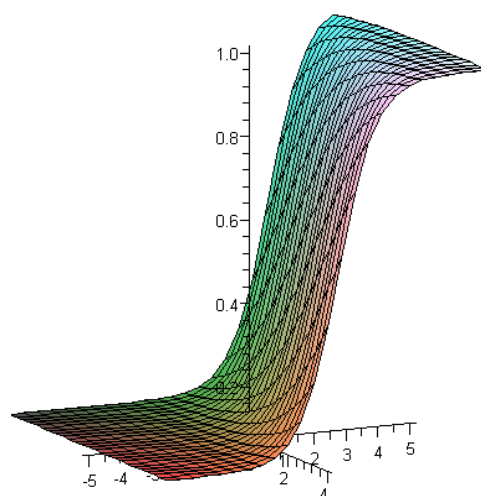
Dengan menggunakan respons peserta tes terhadap butir tes, parameter butir dan parameter kemampuan dapat diestimasi. Ada beberapa software yang dapat digunakan untuk membantu analisis butir dengan pendekatan MIRT. Sampai saat ini, perangkat lunak (*software*) yang ada untuk menganalisis butir dengan MIRT hanyalah untuk model kompensatori. Pada data dikotomi, program yang biasa digunakan adalah MAXLOG, NOHARM dan TESTFACT (Spencer, 2004). NOHARM dapat mengestimasi parameter-parameter model 1,2,3 parameter, namun hanya terbatas pada parameter butir saja, dan belum dapat mengestimasi parameter kemampuan peserta. MAXLOG mengestimasi parameter model kompensatori untuk 2 parameter. TESTFACT merupakan perangkat lunak untuk analisis dengan MIRT baik untuk estimasi parameter butir maupun parameter kemampuan peserta, namun belum dapat mengestimasi parameter  $c$  (*pseudo-guessing*) sekaligus, namun perlu diinputkan setelah melalui analisis dengan software lain, BILOG misalnya.

Sebagai contoh analisis, digunakan perangkat tes Ujian Akhir (UN) mata pelajaran matematika tahun 2005 untuk sekolah menengah pertama (SMP). Setelah melalui analisis factor eksploratori untuk mengetahui banyaknya dimensi, model yang paling cocok yakni model analisis 2 dimensi. Dengan kata lain, UN mata pelajaran matematika mengukur dua dimensi kemampuan, yang kemudian dinamai dengan kemampuan spasial dan kemampuan numerik. Naskah butir soal nomor 29 sebagai berikut.

29. Seorang anak bermain layang-layang dengan panjang benang 80 meter. Jika benang dianggap lurus, sudut antara benang dan tanah  $55^\circ$ . Berapa tinggi layang-layang?

- |             |                        |
|-------------|------------------------|
| a. 45,84 m  | Diketahui :            |
| b. 65,52 m  | Sin $55^\circ = 0,819$ |
| c. 97,68 m  | Cos $55^\circ = 0,573$ |
| d. 114,24 m | Tan $55^\circ = 1,428$ |

Butir ini dapat dijawab dengan benar dengan menggunakan kemampuan spasial dan numerik. Setelah diestimasi, butir ini memiliki parameter tingkat kesulitan sebesar 1,700, *pseudo-guessing* sebesar 0,093, parameter daya pembeda untuk dimensi spasial sebesar 0,347 dan dimensi numeric sebesar 1,025. Kurva karakteristiknya digambarkan sebagai berikut.



Gambar 3. Permukaan karakteristik butir nomor 29 UN matematika SMP 2005

Dalam bidang pendidikan dan psikologi, teori respons butir multidimensi dapat diterapkan untuk mengukur kemampuan umum ataupun kemampuan psikologis peserta tes. Menurut Ackerman, dkk. (2003), teori respons butir multidimensi dapat diarahkan pada pengembangan tes, memperoleh informasi diagnostik tentang estimasi kemampuan, keberfungsian butir diferensial (*differential item functioning, DIF*), dan model teori respons butir untuk data politomous. Segall (2000) memperkuat pendapat ini dengan menyatakan bahwa teori respons butir multidimensi dapat digunakan untuk pemilihan butir, dalam rangka memprediksi pembelajaran maupun mengestimasi kemampuan peserta didik.

## KESIMPULAN

Hubungan antara parameter butir dan parameter peserta dengan peluang menjawab benar dinyatakan dalam persamaan regresi logistik multivariat, yang merupakan salah satu cara analisis butir dengan pendekatan teori respons butir multidimensi. Dengan menggunakan respons peserta tes terhadap butir tes, parameter butir dan parameter kemampuan dapat diestimasi, yang dapat dilakukan dengan bantuan *software* komputer.

## REFERENSI

- Ackerman, T.A., dkk. (2003). Using multidimensional item response theory to evaluate educational and psychological tests. *Educational Measurement*, 22, 37-53.
- Bolt, D.M. & Lall, V.M. (2003). Estimation of compensatory and noncompensatory multidimensional item response models using Markov chain Monte-Carlo. *Applied Psychological Measurement*, 27, 395-414.
- Bryant, D. U. (tth). Directional item information for the multidimensional three-parameter logistik model. *Running head : multidimensional item information*.
- Hambleton, R.K., Swaminathan, H & Rogers, H.J. (1991). *Fundamental of item response theory*. Newbury Park, CA : Sage Publication Inc.
- Hambleton, R.K. & Swaminathan, H. (1985). *Item response theory*. Boston, MA : Kluwer Inc.

- Kirisci, L., Hsu, T., & Yu, L. (2001). Robustness of item parameter estimation programs to assumptions of unidimensionality and normality. *Applied Psychological Measurement*, 25, 146-162.
- Reckase, M.D. (1997). A linear logistic multidimensional model for dichotomous item response data. In W.J. Linden & R.K. Hambleton (Eds), *Handbook of modern item response theory* (pp. 271-286). New York : Springer.
- Spencer, S.G. (2004). The strength of multidimensional item response theory in exploring construct space that is multidimensional and correlated. *Dissertation*. Brigham Young University.
- Spray, J.A., dkk. (1990). Comparison of two logistic multidimensional item response theory models. *ACT Research Report Series*. United States Government.
- Xitao Fan. (1998). Item response theory and classical test theory : an empirical comparison of their item/person statistics. *Educational and Psychological Measurement*, 58, 357-381.