

**MAREK KRYNIEWSKI**

Zespół Szkół Energetycznych w Gdańsku

## **UŻYCIE JĘZYKA PROGRAMOWANIA ICL DO SZACOWANIA PARAMETRÓW KRZYWEJ CHARAKTERYSTYCZNEJ ZADANIA**

### **1. Przypomnienie trójparametrycznego modelu IRT**

Najlepszym modelem do zastosowania w pomiarze dydaktycznym jest model trójparametryczny. Jest on modelem najogólniejszym, sformułowanym po raz pierwszy przez Birnbauma (1968 r.). W modelu tym przyjmuje się, że prawdopodobieństwo udzielenia prawidłowej odpowiedzi na zadanie testowe zależy od trzech parametrów charakteryzujących to zadanie:

- jego mocy różnicującej                      parametr **a**,
- trudności zadania testowego                parametr **b**,
- współczynnika zgadywania                parametr **c**.

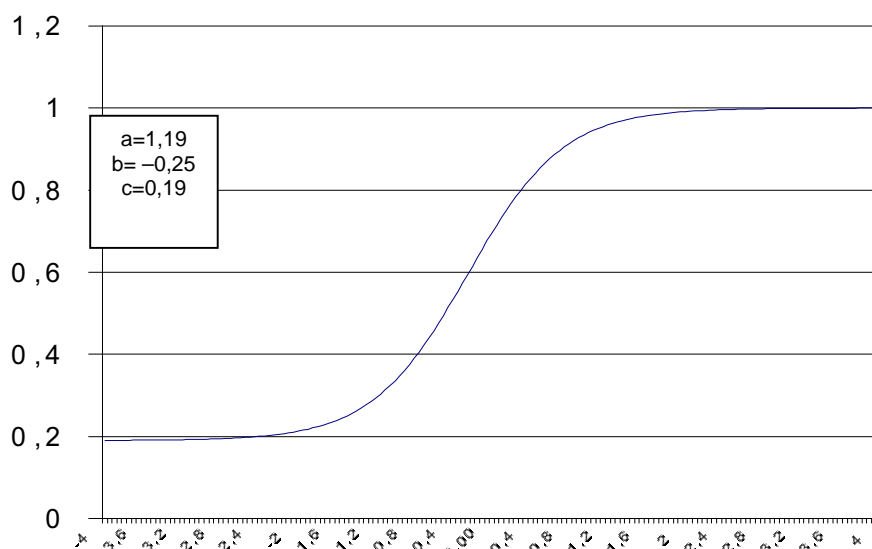
Funkcję opisującą związek między prawdopodobieństwem  $P_i(\theta)$  udzielenia odpowiedzi prawidłowej na i-te zadanie a poziomem wiedzy (umiejętności) w znormalizowanej skali można przedstawić jako:

$$P_i(\theta) = c_i + \frac{1 - c_i}{1 + e^{D \cdot a_i (\theta - b_i)}}, \quad i = 1, 2, \dots, n$$

gdzie:

- n – ilość zadań w teście,
- $c_i$  – współczynnik zgadywania dla i-tego zadania,
- $b_i$  – współczynnik trudności dla i-tego zadania,
- $a_i$  – współczynnik mocy różnicującej dla i-tego zadania,,
- e – liczba Eulera w przybliżeniu  $e=2,718\dots$  jest używana jako podstawa logarytmu naturalnego,
- D – stała maksymalizująca dopasowanie krzywej logistycznej do *ogivy* rozkładu normalnego;  $D=1,7$ .

Najczęściej bowiem, aby rzetelność obliczeń podczas estymacji współczynników **a**, **b**, **c** była zadowalająca, trzeba przetworzyć dane uzyskane z przebadania przynajmniej 1000 osób.



Rys. 1. Przykład krzywej charakterystycznej dla określonych parametrów

## 2. Użycie języka programowania ICL

### Dostęp do języka programowania ICL

Wiele osób zajmujących się pomiarem dydaktycznym chętnie obok klasycznych miar zastosowałoby miary probabilistyczne. Konieczne więc staje się posiadanie programów komputerowych, które umożliwią obliczenie miar probabilistycznych. Dostępne są komercyjne programy, których cena nie jest przystępna dla osoby chcącej podjąć pierwsze próby obliczeń z użyciem teorii probabilistycznej. Istnieje jednak darmowe i proste narzędzie, które może być zastosowane przez osobę zajmującą się pomiarem. Jest nim język programowania ICL. W celu korzystania z tego narzędzia konieczne jest ściągnięcie z Internetu pliku **icl\_win.zip**.

Plik **icl\_win.zip** zawiera:

- **icl.exe** – program obliczeniowy,
- **icl\_manual.pdf** – podręcznik użytkownika,
- **release\_notes.txt** – aktualną notatkę o wersji ICL,
- **examples** – folder z przykładami.

Plik ten był udostępniony pod adresami: [www.b-a-h.com/software/irt/icl/](http://www.b-a-h.com/software/irt/icl/) oraz [mirror.optusnet.com.au/sourceforge/s/ss/ssm/](http://mirror.optusnet.com.au/sourceforge/s/ss/ssm/)

Po rozpakowaniu pliku **icl\_win.zip** program **icl.exe** jest gotowy do użycia. Istnieją również wersje dla systemów operacyjnych Mac oraz Linux.

## Podstawy pracy z językiem ICL

Program może pracować w dwóch trybach:

1. **tryb komend** – po uruchomieniu programu icl.exe uruchamiane jest okienko przypominające okienko DOS ze znakiem zachęty %; teraz możemy wydawać komendy,
2. **tryb wsadowy**, który uruchamiany jest poprzez wydanie komendy: icl.exe *plik\_komend.tcl*. Wygodne jest napisanie pliku typu BAT (plik taki tworzymy np. w notatniku jako zwykły plik tekstowy, należy tylko pamiętać, aby w nazwie jako rozszerzenie nadać BAT), np. uruchom.bat, a w nim tylko jedna linijka, np. icl.exe oblicz.tcl, gdzie oblicz.tcl to plik, który zawiera komendy języka ICL. (Osobiście polecam tryb wsadowy. W celu skorzystania z tego języka należy napisać plik BAT oraz plik typu tcl, też jako plik tekstowy. Plik typu BAT można uruchomić poprzez podanie jego nazwy w okienku Uruchom, które jest osiągalne po kliknięciu przycisku Start).

## Podstawy języka ICL

Podstawy języka wytłumaczę na podstawie treści pliku typu tcl.

W treści pliku typu tcl mogą znajdować się linie zaczynające się od znaku #. Są to komentarze ignorowane przez program. Mogą istnieć również puste linie, które są również ignorowane. Wytłumaczenia poszczególnych komend języka ICL znajdują się w treści pliku typu tcl znajdującego się poniżej. Wytłumaczenia te są wykonane jako komentarze po znaku #.

### Treść pliku typu tcl o nazwie artykul.tcl

```
#
# tresc pliku  artykul.tcl
#
# Szacowanie parametrow a,b,c dla modelu IRT
# dla 30 zadań wielokrotnego wyboru
# testu konkursu informatycznego
# dla uczniow gimnazjow
# Gdansk (2003)

# instrukcja okreslenia pliku wyjsciowego
# nazwa pliku artykul.log

output -log_file artykul.log

# dla 30 zadan WW beda szacowane parametry a,b,c

allocate_items_dist 30

# Wczytanie odpowiedzi na zadania WW z pliku artykul.dat
# Kazdy wiersz zawiera odpowiedzi jednego ucznia
# 30 zadan z konkursu sa w kolumnach od 1 do 30
```

```
read_examinees artykul.dat 30i1

# Start szacowania parametrow a,b,c dla wszystkich zadan WW

starting_values_dichotomous

# wykonanie EM iteracji dla obliczen szacowanych parametrow a,b,c
# maksymalnie 50 EM iteracji

EM_steps -max_iter 50

# zapisanie parametrow a,b,c oraz
# rozkladu badanej cechy
# (wiedzy-->niewidocznej[latent], odseparowanej[discrete])
print -item_param -latent_dist

# zakonczenie procesu szacowania
release_items_dist
```

---

### Treść pliku wynikowego typu log o nazwie artykul.log

Poniżej przedstawiony jest plik, który jest wynikiem działania programu napisanego z użyciem języka ICL

IRT Command Language (ICL)  
Version 0.020301

Jul 11, 2005 13:19

Command file artykul.tcl

-----  
*tutaj jest treść pliku typu tcl*  
-----

Number of items: 30  
Number of latent variable points: 40  
Number of examinee groups: 1

Default prior for a-parameters:  
beta a: 1.750 b: 3.000 lower limit: 0.000 upper limit: 3.000  
Default prior for b-parameters:  
beta a: 1.010 b: 1.010 lower limit: -6.000 upper limit: 6.000  
Default prior for c-parameters:  
beta a: 3.500 b: 4.000 lower limit: 0.000 upper limit: 0.500

Read 392 examinee records from file artykul.dat

EM iterations  
(iteration: parameter criterion, marginal posterior mode)

1: 0.366067 -6739.6733  
 2: 0.141608 -6737.7174

.....  
 tutaj są wyniki kolejnych błędów iteracji

.....  
 15: 0.001033 -6736.8092  
 16: 0.000836 -6736.8085

Item Parameter Estimates

(a, b, c for 3PL, 2PL, 1PL; a, b1, b2, ... for GPCM, PCM)

1	0.711500	0.650589	0.207228
2	1.103929	4.122096	0.170985
3	0.486210	-3.846867	0.228719

.....  
 współczynniki dostępne w tabeli 1

27	0.920212	0.832028	0.177283
28	1.200661	2.148231	0.242168
29	0.394483	-2.182686	0.226173
30	0.857350	0.403236	0.204076

Discrete Latent Variable Distributions

(point, probability for group 1, 2, etc)

-4.000000	2.745344e-005
-3.794872	6.106663e-005
-3.589744	1.302378e-004
-3.384615	2.663153e-004

.....  
 3.589744 1.302378e-004  
 3.794872 6.106663e-005  
 4.000000 2.745344e-005

**Treść pliku danych typu dat o nazwie artykul.dat**

Pojedynczy wiersz to odpowiedzi jednego ucznia: 0 – zła odpowiedź, 1 – poprawna odpowiedź. W kolumnach znajduje się zapis poszczególnych zadań testowych. Plik ten jest plikiem tekstowym, może być wykonany w notatniku.

001101110100010000011011101010  
 001101010100010001100111101010

.....  
 pozostałe dane:

.....  
 011111111100010101000111011011  
 101111110101101011010111101011

### 3. Użycie języka ICL dla porównania parametrów klasycznych oraz IRT

#### Zestawienie parametrów klasycznych oraz IRT

Interesujące wydaje się zestawienie klasycznych parametrów testu z parametrami IRT. Poniższe zestawienie wykonano dla testu, który był częścią konkursu informatycznego dla uczniów gimnazjów przeprowadzonego w województwie pomorskim w 2003 r. Test zawierał 30 zadań WW z czterema możliwymi do wyboru odpowiedziami. Uczestniczyło w nim 392 uczniów.

Tabela 1. Zestawienie parametrów klasycznych oraz IRT dla przykładowego testu

numer zadania	różnicowanie		trudność		współczynnik zgadywania
	a-IRT	klasyczna	b-IRT	klasyczna	c_IRT
1	0,712	0,388	0,651	0,515	0,207
2	1,104	0,000	4,122	0,829	0,171
3	0,486	0,154	-3,847	0,041	0,229
4	1,097	0,207	1,767	0,510	0,414
5	0,777	0,358	1,692	0,696	0,182
6	0,738	0,409	-1,019	0,214	0,214
7	0,530	0,296	1,930	0,666	0,188
8	0,642	0,359	-1,172	0,202	0,229
9	0,096	0,133	0,967	0,408	0,242
10	1,188	0,469	-0,248	0,347	0,188
11	1,159	0,434	1,401	0,763	0,111
12	1,562	0,336	1,638	0,811	0,116
13	1,286	0,373	1,156	0,615	0,247
14	0,620	0,310	-1,406	0,176	0,228
15	0,652	0,220	2,775	0,732	0,214
16	0,948	0,321	1,679	0,714	0,184
17	0,773	0,425	-0,150	0,372	0,199
18	0,467	0,342	-0,222	0,355	0,232
19	0,622	0,337	1,664	0,648	0,203
20	1,034	0,409	1,523	0,773	0,107
21	0,747	0,433	1,164	0,666	0,126
22	0,528	0,316	0,199	0,408	0,241
23	0,993	0,438	-0,716	0,242	0,214
24	0,526	0,282	-1,782	0,156	0,227
25	0,197	0,187	1,355	0,457	0,249
26	0,475	0,229	2,560	0,689	0,204
27	0,920	0,433	0,832	0,589	0,177
28	1,201	0,181	2,148	0,717	0,242
29	0,394	0,255	-2,183	0,161	0,226
30	0,857	0,424	0,403	0,482	0,204

## Wnioski wynikające z zestawienia parametrów klasycznych oraz IRT

### Współczynnik a: różnicowanie

Istnieje słaba zależność pomiędzy klasyczną miarą różnicowania a miarą IRT (współczynnik korelacji Persony między tymi wielkościami wynosi 0,302). Są to wielkości mierzące inne wartości. W teorii klasycznej korelacyjny współczynnik mocy różnicującej przedstawia siłę związku w całym zbiorze wyników, a największy wpływ na jego wielkość mają wyniki najwyższe i najniższe, natomiast w teorii probabilistycznej współczynnik  $a$  pokazuje nachylenie krzywej w punkcie największego zróżnicowania, a więc największy wpływ na jego wielkość mają zwykle wyniki środkowe.

### Współczynnik b: trudność

Istnieje ścisły związek pomiędzy klasyczną miarą trudności a miarą IRT (współczynnik korelacji Persony między tymi wielkościami wynosi 0,923). Są to wielkości mierzące ten sam parametr testu, dla teorii klasycznej oraz IRT.

### Współczynnik c: współczynnik zgadywania

Średni współczynnik zgadywania wynosi 0,207. Jest wielkością zbliżoną do prawdopodobieństwa zgadywania, która dla tego testu wynosi 0,25 ( $1/p$  gdzie  $p$  – ilość możliwych odpowiedzi na zadanie WW,  $p = 4$ ). Maksymalny współczynnik zgadywania to 0,414, a minimalny – 0,107.