

ПЕД	
	измерения

РАСЧЁТ ИСТИННОЙ ОЦЕНКИ ИСПЫТУЕМЫХ, ПОСТРОЕНИЕ ГРАФИЧЕСКИХ ОБРАЗОВ ТЕСТА И ЗАДАНИЙ ПОСРЕДСТВОМ EXCEL

Роман Дубинка

г. Барановичи, Республика Беларусь
xxxroman@bk.ru

В данной статье рассмотрено понятие истинной оценки, способ ее расчета вручную и при помощи табличного процессора Microsoft Excel 2007; построение кривой характеристики задания и ее анализ. Приведены примеры, а также упражнения для закрепления теоретического и практического материала.

Ключевые слова: истинная оценка, графический образ теста, уровень подготовленности испытуемых, вероятность правильного ответа.

Понятие истинной оценки. Определение и построение графического образа теста

В математической теории измерений (Item Response Theory) рассматривается как тест в целом, так и каждое его задание в отдельности. Чтобы оценить результат выполнения теста используются различные шкалы. Для данных, представленных в дихотомической шкале, правильный ответ на задание дает оценку 1, неправильный — 0. Сумма полученных по каждому заданию оценок будет являться исходной оценкой теста. Эта оценка будет всегда являться целым числом и колебаться от 0 до N в зависимости от числа заданий в тесте. Если испытуемый получит этот тест вновь и не вспомнит, как он ответил в предыдущий раз на задания, будет получена другая оценка. Гипотетически, испытуемый может выполнять тест много раз и получать ряд различных оценок. Можно предположить, что все оценки будут варьировать вокруг какой-то средней величины.

В математической теории измерений используется формула расчёта истинной оценки уровня подготовленности, предложенная D.N. Lawley)¹.

Формула для расчёта истинной оценки уровня подготовленности испытуемого i :

$$TS_i(\theta_i) = \sum_{i=1}^K P(\theta_i)$$

где $TS_i(\theta_i)$ — истинная оценка для испытуемых с уровнем способности θ_i . Бейкер обозначает задание $P(\theta_i)$ в зависимости от модели построения кривой характеристики каждого отдельного задания.

Для наглядности рассмотрим пример, в котором рассчитаем истинную оценку теста, состоящего из 7 заданий (см. табл. 1). В этом примере уровень подготовленности испытуемого условно принимается равным одному логиту ($\theta_i = 1$).

Таблица 1.

Номер задания	Параметр крутизны графика задания a_j	Параметр b
1	0,73	0,18
2	0,56	0,2
3	1	0,6
4	0,98	0,01
5	0,851	0,4
6	0,76	0,3

Произведем расчеты вероятностей с помощью формулы для двухфакторной модели.

$$P(\theta) = \frac{1}{1 + e^{-1,7a(\theta - \beta)}}$$

Задание 1

$$P_1(\theta) = \frac{1}{1 + e^{-1,7 \cdot 0,73(\theta - (-0,18))}} = 0,7343508745.$$

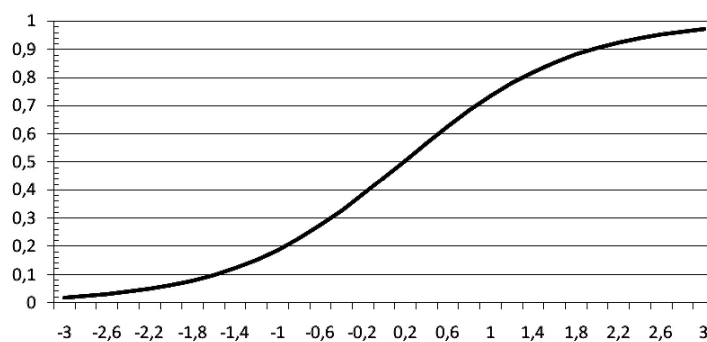


Рис. 1

ПЕД
измерения

Задание 2

$$P_2(\theta) = \frac{1}{1 + e^{-1.7056(\theta - 0.2)}} = 0,68170101.$$

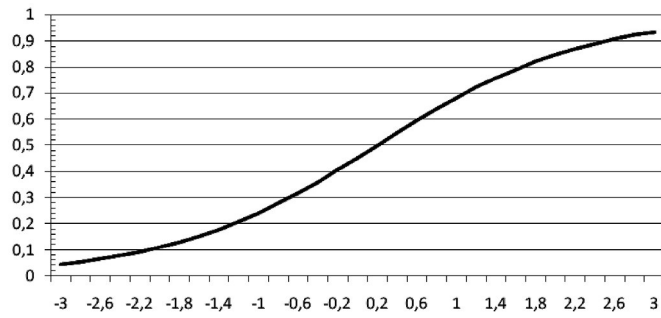


Рис. 2

Задание 3

$$P_3(\theta) = \frac{1}{1 + e^{-1.74(\theta - 0.6)}} = 0,663738697.$$

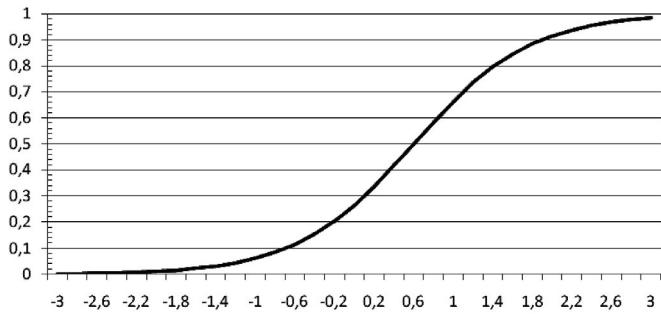


Рис. 3

Задание 4

$$P_4(\theta) = \frac{1}{1 + e^{-1.7098(\theta - 0.01)}} = 0,83880183.$$

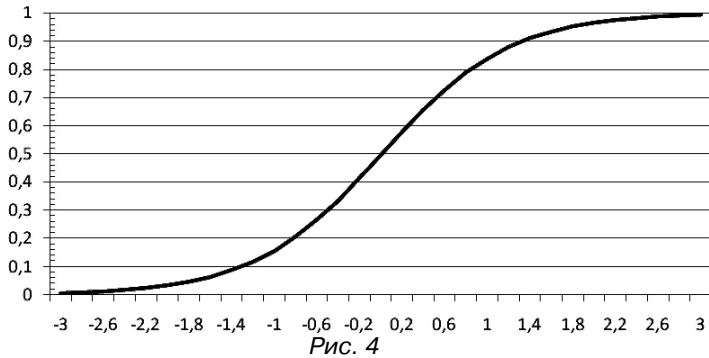
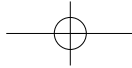


Рис. 4



Задание 5

$$P_5(\theta) = \frac{1}{1 + e^{-1.70,851(\theta - 0.4)}} = 0,704333534.$$

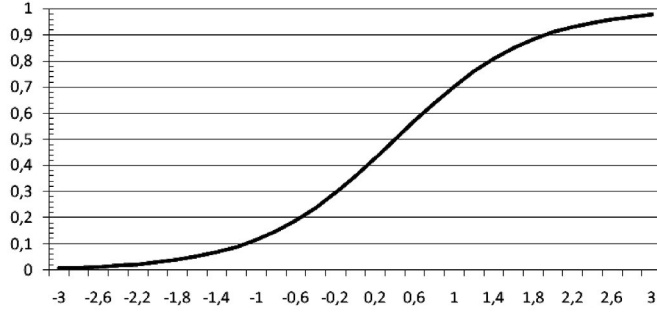


Рис. 5

Задание 6

$$P_6(\theta) = \frac{1}{1 + e^{-1.70,76(\theta - 0.8)}} = 0,711852864.$$

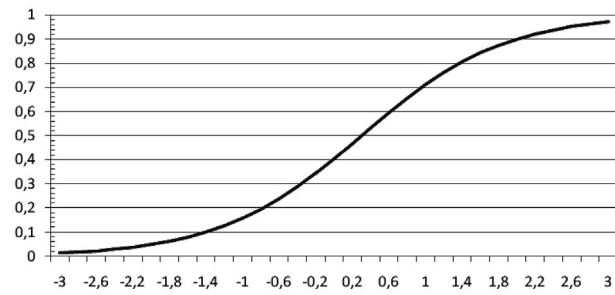


Рис. 6

Задание 7

$$P_7(\theta) = \frac{1}{1 + e^{-1.70,62(\theta - 0.02)}} = 0,737481607.$$

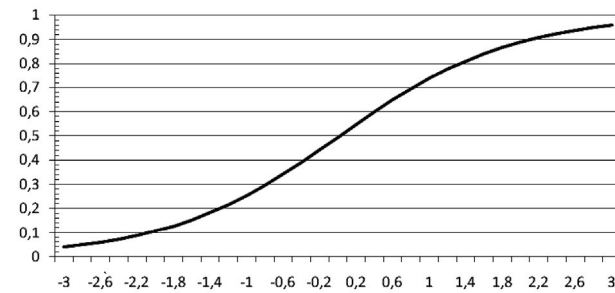
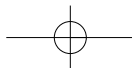
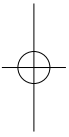
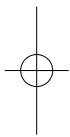


Рис. 7

Теория



ПЕД
измерения

Теперь, чтобы получить истинную оценку при $\theta = 1$, вероятности правильного ответа на все задания «теста» суммируем.

$$TS_j = 5,072418.$$

Таким образом, испытуемые с уровнем подготовленности, равным один логит, имеют наиболее вероятную оценку 5,072418. После округления до целого числа это будет тестовый балл 5. Причем данную оценку может не получить ни один из испытуемых, т.к. она — теоретическое среднее число всех вероятных оценок, которые могут получить испытуемые с уровнем способности, равным 1 логит, выполнявшими тест. Фактически тест должен содержать гораздо больше заданий, однако предполагает-

ся, что истинная оценка должна оставаться прежней.

Расчёты, производимые выше, были только для одной точки на шкале подготовленности. (Аналогичные расчёты могут быть выполнены для любой точки от $-\infty$ до ∞ .) Соединив совокупность этих точек можно получить график функции теста от уровня подготовленности. Вертикальная ось — ожидаемая истинная оценка, которая может колебаться от 0 до количества заданий в тесте. Горизонтальная ось — шкала уровня подготовленности испытуемых. Этот график, выполненный в виде плавно восходящей линии, и будет характеризовать тест. Ниже приведен пример графика теста, состоящего из 10 заданий.

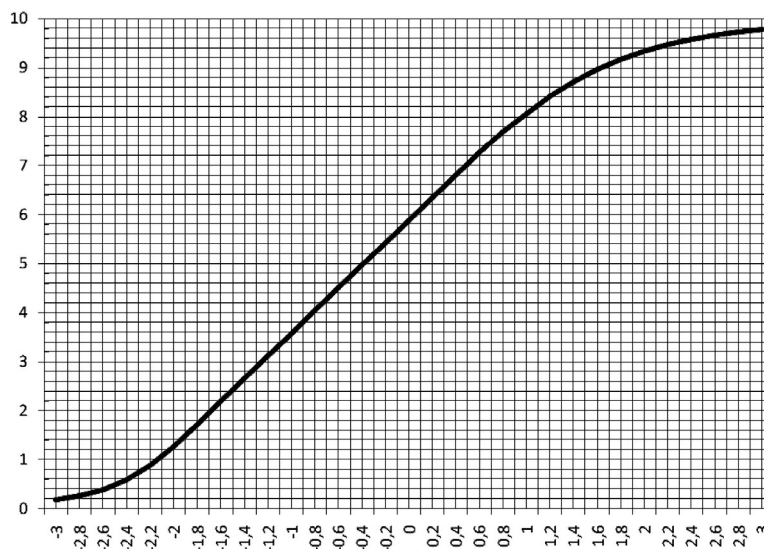


Рис. 8

График теста является функциональной зависимостью между уровнем подготовленности и вероятной истинной оценкой испытуемого. При помощи данного графика можно найти истинную оценку испытуемых на любом уровне подготовленности.

При использовании одной двухпараметрической модели измерения левый конец графика задания будет стремиться к нулю, позволяя таким образом оценить самый низкий уровень подготовленности; правый конец графика будет стремиться к количеству заданий в тесте k . При использовании трехпараметрической модели левый конец графика будет стремиться не к нулю, а к априорной оценке, которую можно получить по вероятности угадывания правильного ответа в каждом задании.

Расчеты в Microsoft Excel

Рассчитаем истинные оценки для данных из *таблицы 1*. Для начала работы запускаем Microsoft Office: Меню «Пуск»\Программы (Все программы)\ Microsoft Office\Microsoft Office Excel

В открывшемся окне рабочей **Книги 1** в ячейке первой строки вводим заголовки таблицы: *Задание (№ задания)*; a (*фактор a*), b (*фактор b*).

Заполняем таблицу данными.

	A	B	C	D
1	Задание	a	b	
2	1	0,73	0,18	
3	2	0,56	0,2	
4	3	1	0,6	
5	4	0,98	0,01	
6	5	0,851	0,4	
7	6	0,76	0,3	
8				

Помещаем курсор в ячейку $F1$ и приступаем к заполнению уровней способности $\theta \in (-3; 3)$ направо по горизонтали с шагом 0,2 — это необходимо для того, чтобы кривая характеристики теста получилась плавной (S-образной). Чтобы ускорить процесс заполнения ячеек диапазона $F1:AJ1$, можно использовать формулу ($=F1+0,2$), которая вводится в ячейку $G1$ после ввода первого значения уровня способности в ячейку $F1$.

	A	B	C	D	E	F	G
1	Задание	a	b			-3	$=F1+0,2$
2	1	0,73	0,18				
3	2	0,56	0,2				
4	3	1	0,6				
5	4	0,98	0,01				
6	5	0,851	0,4				
7	6	0,76	0,3				
8							

Содержимое ячейки $G1$ копируется во все ячейки диапазона $F1:AJ1$ при помощи ее растягивания (перемещаем курсор к правому нижнему углу ячейки $G1$, когда он примет вид черного крестика, зажимаем левую клавишу мыши и перемещаем курсор вправо пока не будет выделен необходимый диапазон ячеек).

Теория

Книги

ПЕД
измерения

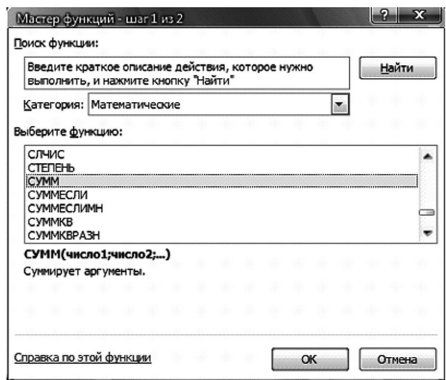
Перемещаем курсор в ячейку F2 для расчета вероятности правильного ответа. Вводим формулу двухфакторной модели $\ll=1/(1+(EXP(-1,7*SB2*(FS1-SC2))))\gg^1$

СУММ $\ll=1/(1+(EXP(-1,7*SB2*(FS1-SC2))))\gg$					
	A	B	C	D	E
1	Задание	a	b		F
2	1	0,73	0,18		$\ll=1/(1+(EXP(-1,7*SB2*(FS1-SC2))))\gg$
3	2	0,56	0,2		
4	3	1	0,6		
5	4	0,98	0,01		
6	5	0,851	0,4		
7	6	0,76	0,3		
8					

Копируем формулу растягиванием сначала по вертикали до ячейки F7, выделяем диапазон ячеек F2:F7 и растягиваем его по горизонтали до столбца AJ включительно. В результате таблица примет вид:

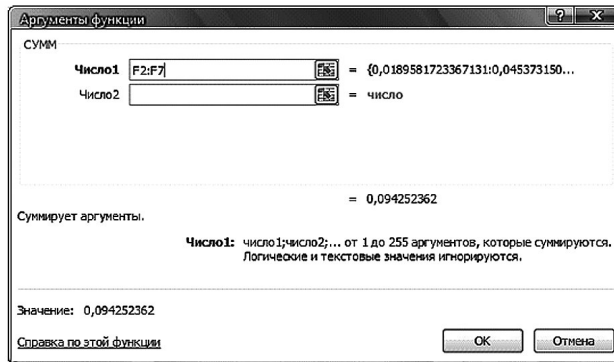
AJ7 $\ll=1/(1+(EXP(-1,7*SB7*(AJ51-SC7))))\gg$																										
1	Задание	a	b																							
2	1	0,73	0,18																							
3	2	0,56	0,2																							
4	3	1	0,6																							
5	4	0,98	0,01																							
6	5	0,851	0,4																							
7	6	0,76	0,3																							

При помощи функции СУММ рассчитаем сумму вероятностей для каждого уровня способности. Отступаем одну строку от массива с вероятностями и помещаем курсор в ячейку F9, щелкаем по значку $\ll=f_x\gg$.
В открывшемся диалоговом окне из раскрывающегося списка Категория выбираем Математические, а в появившемся списке ниже указываем СУММ.



Щелкаем по кнопке **Ok** и в поле **Число 1** вводим диапазон ячеек F2:F7 либо, щелкнув по кнопке, выделяем необходимый диапазон в рабочей книге.

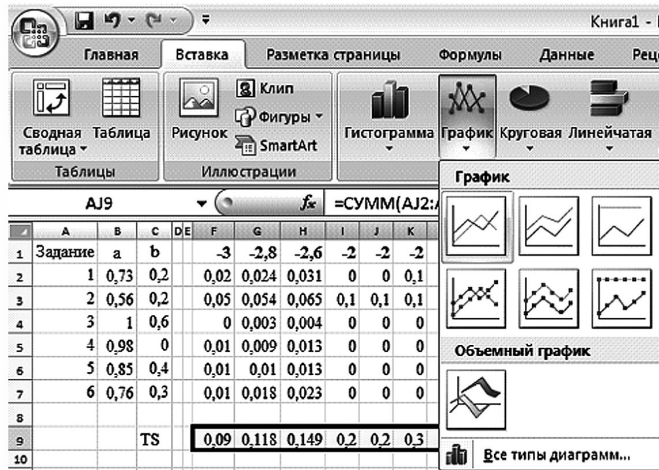
1
Дубинка Р.В.
Построение графика информационной функции посредством Excel. ПИ. 2009. № 1.



Введенную формулу растягиваем на диапазон ячеек F9:AJ9. В результате получаем значения исходных оценок.

		=СУММ(AJ2:AJ7)										
A	B	C	D	E	F	G	H	I	J	K	L	
1	Задание	a	b		-3	-2,8	-2,6	-2	-2	-2		
2		1	0,73	0,2	0,02	0,024	0,031	0	0	0,1		
3		2	0,56	0,2	0,05	0,054	0,065	0,1	0,1	0,1		
4		3	1	0,6	0	0,003	0,004	0	0	0		
5		4	0,98	0	0,01	0,009	0,013	0	0	0		
6		5	0,85	0,4	0,01	0,01	0,013	0	0	0		
7		6	0,76	0,3	0,01	0,018	0,023	0	0	0		
8												
9				TS	0,09	0,118	0,149	0,2	0,2	0,3		
10												

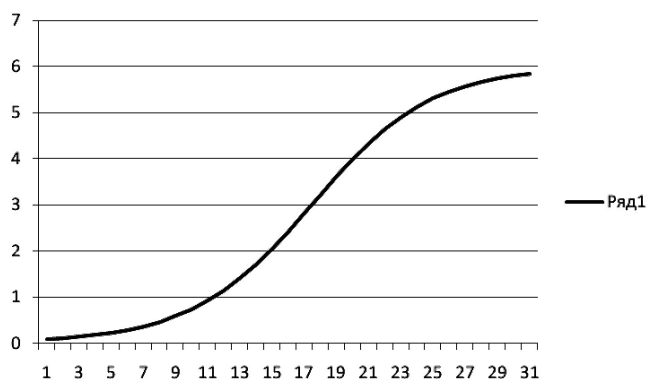
Для построения кривой характеристики теста выделяем диапазон ячеек с истинными оценками и на вкладке меню Вставка щелкаем по кнопке График.



В результате получим график, имеющий следующий вид:

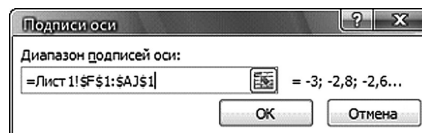
Теория

ПЕД
измерения



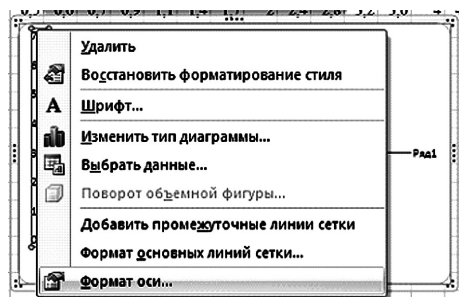
Однако он не совсем корректно отображает подписи к осям, поэтому нуждается в небольшой коррекции. Для начала следует удалить область с надписью **Ряд**. Щелкаем по данной области левой клавишей мыши и, если она выделена, нажимаем клавишу *delete*. Далее необходимо вместо цифр 1,2,3, ..., 31 установить значения уровней способности -3, -2,8, -2,6, ..., 3. Щелкаем по вкладке меню **Конструктор** (область **Работа с диаграммами**), а

затем по кнопке **Выбрать данные** . В правой части открывшегося диалогового окна **Подписи горизонтальной оси (Категории)** щелкаем по кнопке **Изменить** и выделяем диапазон ячеек **F1:AJ1**.

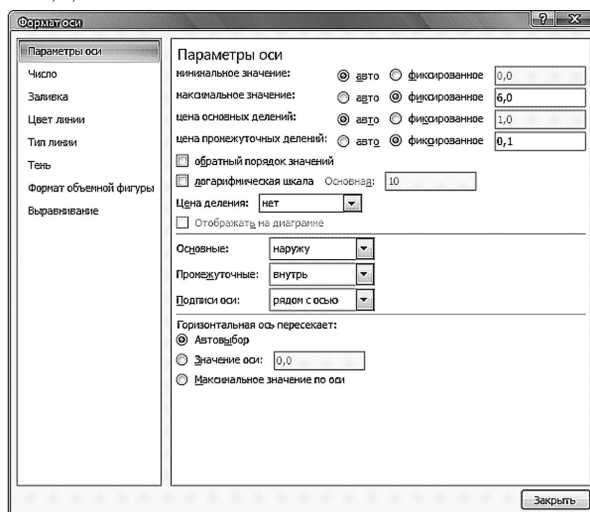


Для завершения щелкаем по кнопке **Ок** в окне **Подписи оси** и в окне **Выбор источника данных**.

Теперь необходимо изменить ось значений **ОУ**, т.к. в тесте всего 6 заданий (это важно в случаях, когда нет исходных данных и необходимо определить количество заданий в тесте). Щелкаем правой клавишей мыши по области значений и из раскрывающегося списка меню выбираем **Формат оси**.



В разделе **Параметры оси** устанавливаем **Максимальное значение**: фиксированное 6,0 и цена промежуточных делений: фиксированное 0,1, как показано ниже.



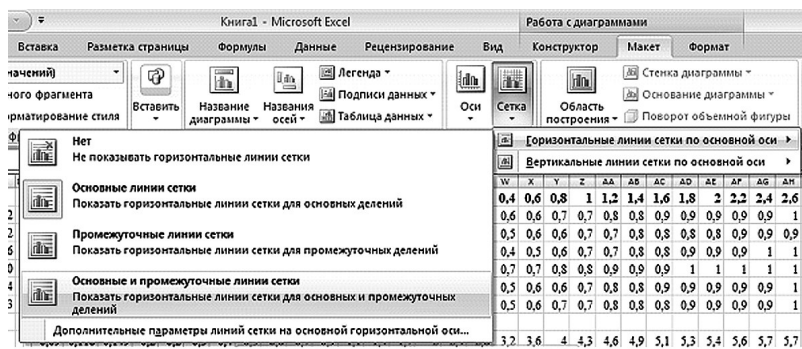
Для завершения щелкаем по кнопке **Заккрыть**.

При построении графика **Excel** автоматически устанавливает отображение только основных горизонтальных линий сетки. Для анализа кривой характеристики теста их недостаточно, поэтому необходимо добавить промежуточные горизонтальные линии, а также основные и промежуточные вертикальные линии сетки. Вкладка меню *Макет \ Сетка \ Горизонтальные (Вертикальные) линии сетки по основной оси \ Основные и промежуточные линии сетки*.

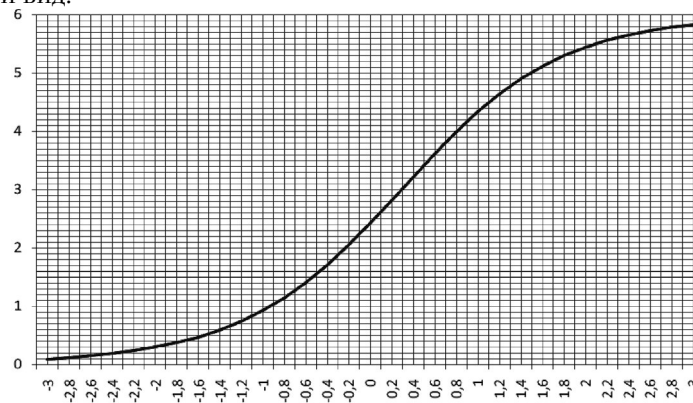
Теория

12/09/09

ПЕД
измерения

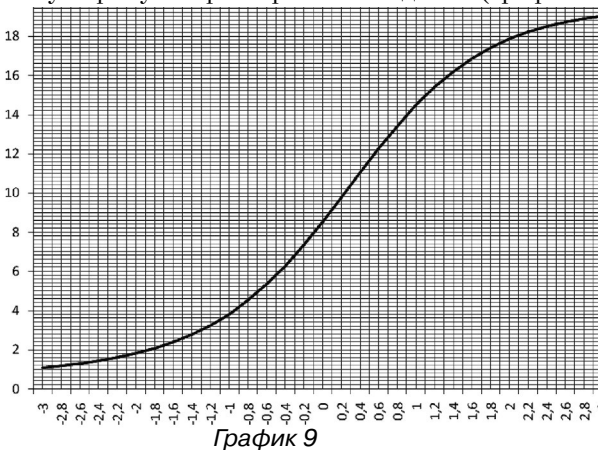


После указанных выше действий график примет окончательный вид.



Практические упражнения

1. Используя кривую характеристики задания (график 9), определить:



а) количество заданий в тесте, для которого построена кривая, если при расчетах использовалась однопараметрическая модель.

б) минимальную и максимальную истинную оценку

в) значение истинной оценки при уровне способности $\theta = 0$.

2. По данным из табл. 2 рассчитать значения истинных оценок для целых значений $\theta \in (-5; 5)$. Построить график функции теста и определить:

Таблица 2

Задание	b
1	0,2
2	0,1
3	0,76
4	0,01
5	0
6	0,8
7	0,2
8	0,4
9	0,12
10	0,02

а) истинную оценку, которую должны получить испытуемые с уровнем подготовленности -5 логитов;

б) уровень подготовленности, которому соответствует оценка 9,9.

3. Для данных из табл. 3 при помощи трехфакторной модели рассчитать значения истинных оценок для целых значений $\theta \in (-10; 10)$. Построить графический образ теста.

Таблица 3.

Задание	b	a	c
1	0,2	0,63	0,23
2	0,1	0,9	0,2
3	0,76	0,2	0,1
4	0,01	0,9	0,61
5	0	0	0,1
6	0,8	1,22	0
7	0,2	-0,56	0,24
8	0,4	0,2	0,03
9	0,12	0,23	0,32
10	0,02	0,46	0,02

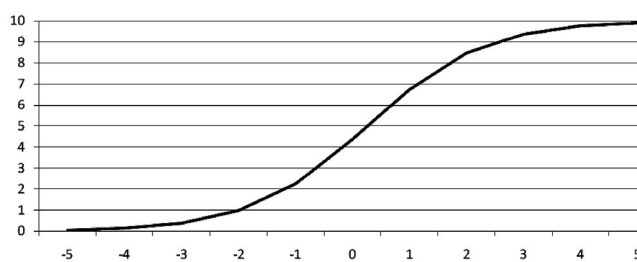
Ответы на упражнения:

1. а) $N = 20$;
б) $TS_{\min} \approx 1$; $TS_{\max} \approx 19$;
в) 8,7

2.

-5	-4	-3	-2	-1
0,0535	0,144	0,38	0,976	2,246

0	1	2	3	4	5
4,368	6,74	8,469	9,3713	9,758	9,91



а) 0,0535;

б) $\theta = 5$.

Теория

Задания

ПЕД
измерения

3.

-10	-9	-8	-7	-6	-5	-4
3,12318	3,14905	3,18525	3,23571	3,30574	3,40265	3,5373

-3	-2	-1	0
3,72857	4,01907	4,52225	5,44527

1	2	3	4	5	6	7
6,71284	7,56749	7,95577	8,18373	8,34506	8,46481	8,55393

8	9	10
8,6197	8,66779	8,70265

