



НФПК



<http://ictlit.com>

Систематический подход к разработке тестовых заданий

**Применение Байесовских сетей для
анализа результатов**

Г. М. Васин

План доклада

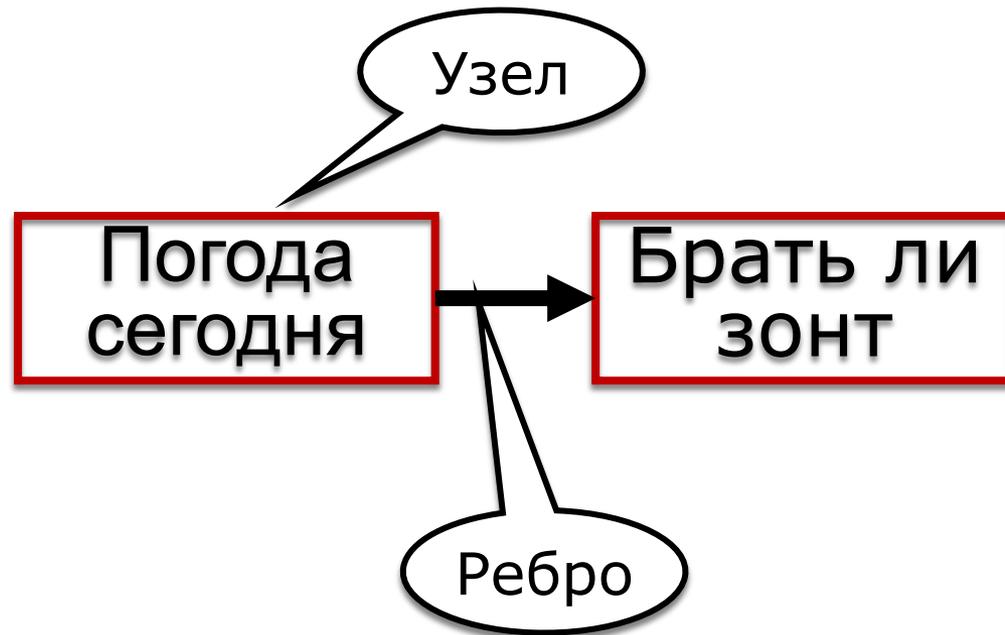
- ❖ Что такое сети Байеса?
- ❖ Несколько примеров
- ❖ Условная вероятность и ее выражение в таблицах
- ❖ Теорема Байеса
- ❖ Метод функциональной тэты, применение модели Samejima
- ❖ Преимущества Байесовских сетей по сравнению с IRT
- ❖ Контроль качества результатов



Байесовская сеть - направленный ациклический граф



Пример сети



Простой пример

Weather					
Sunny	25.0	██			
Rainy	25.0	██			
Cloudy	25.0	██			
Snowy	25.0	██			

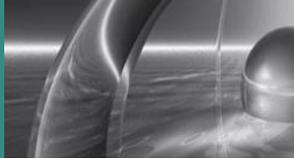
Weather	Umbrella	
	yes	no
sunny	10%	90%
rainy	90%	10%
cloudy	50%	50%
snowy	20%	80%

Umbrella					
Yes	42.5	██			
No	57.5	██			

$$(.10)(.25) + (.90)(.25) + (.50)(.25) + (.20)(.25) = .425$$

$$(.90)(.25) + (.10)(.25) + (.50)(.25) + (.80)(.25) = .575$$

$$P(u = x_i) = \sum_W p(u = x_i | w_j) * p(w_j)$$



Простой пример

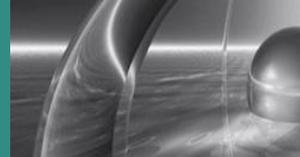
Weather				
Sunny	0			
Rainy	100			
Cloudy	0			
Snowy	0			

Weather	Umbrella	
	yes	no
sunny	10%	90%
rainy	90%	10%
cloudy	50%	50%
snowy	20%	80%

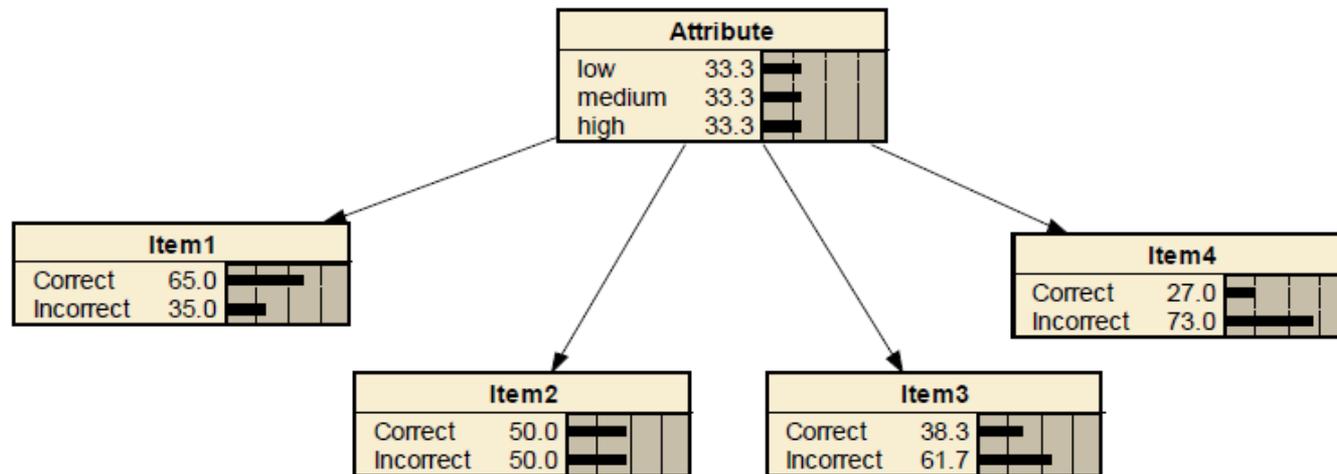
Umbrella				
Yes	90.0			
No	10.0			

$$P(u = x_i) = \sum_W p(u = x_i | w_j) * p(w_j)$$

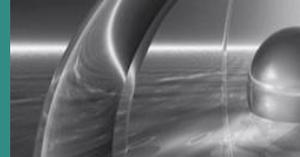
Обновление сети



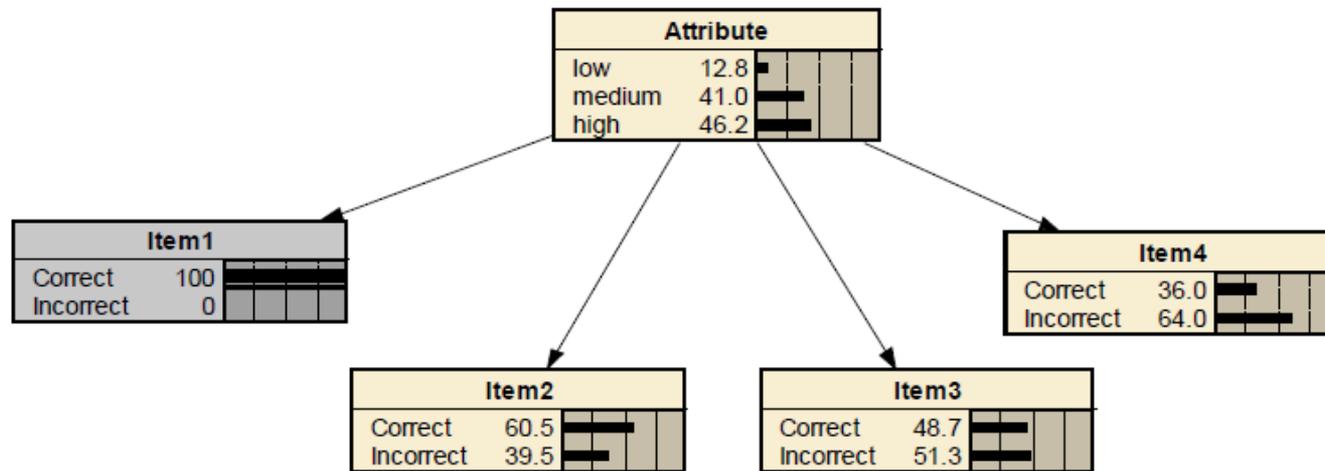
Attribute	Item 1		Item 2		Item 3		Item 4	
	correct	incorrect	correct	incorrect	correct	incorrect	correct	incorrect
low	25%	75%	20%	80%	10%	90%	1%	99%
medium	80%	20%	40%	60%	20%	80%	20%	80%
high	90%	10%	90%	10%	85%	15%	60%	40%



Обновление сети



Attribute	Item 1		Item 2		Item 3		Item 4	
	correct	incorrect	correct	incorrect	correct	incorrect	correct	incorrect
low	25%	75%	20%	80%	10%	90%	1%	99%
medium	80%	20%	40%	60%	20%	80%	20%	80%
high	90%	10%	90%	10%	85%	15%	60%	40%



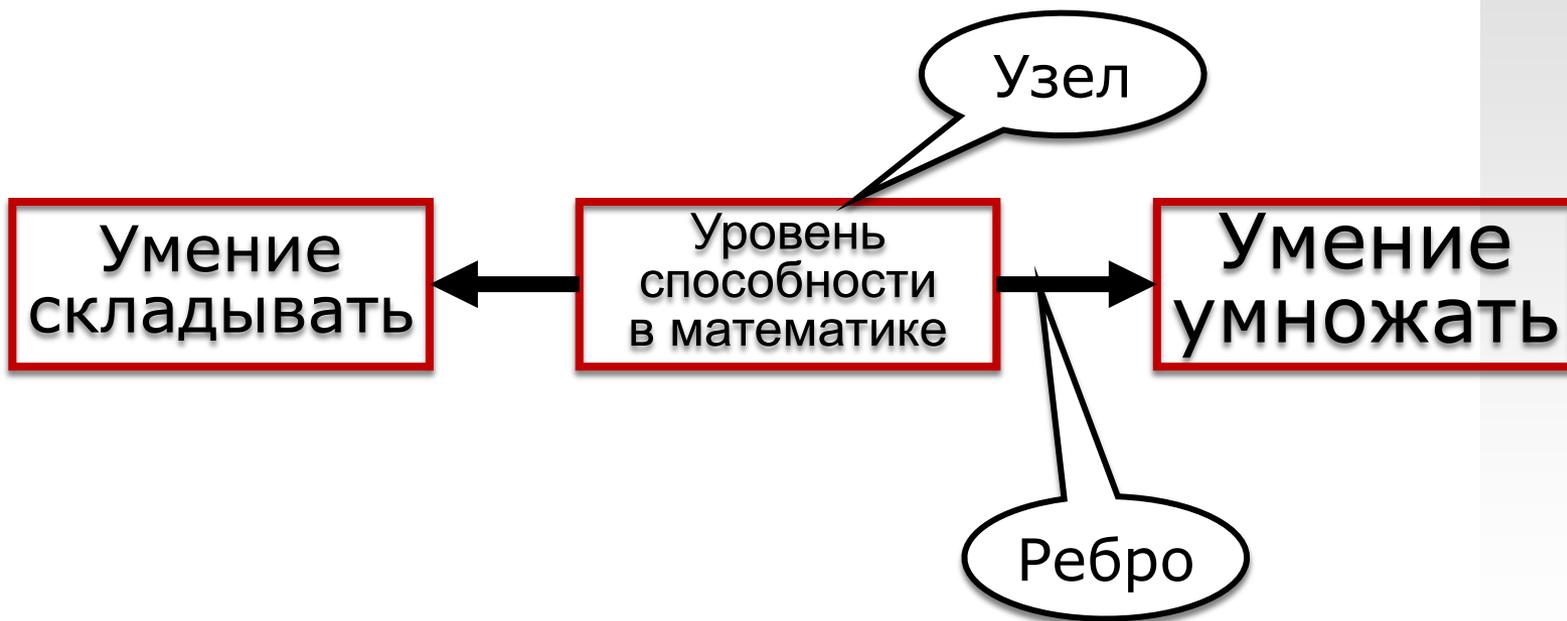
Теорема Байеса

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$

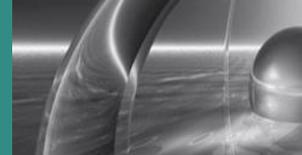


Вероятность
А при услов
ии В

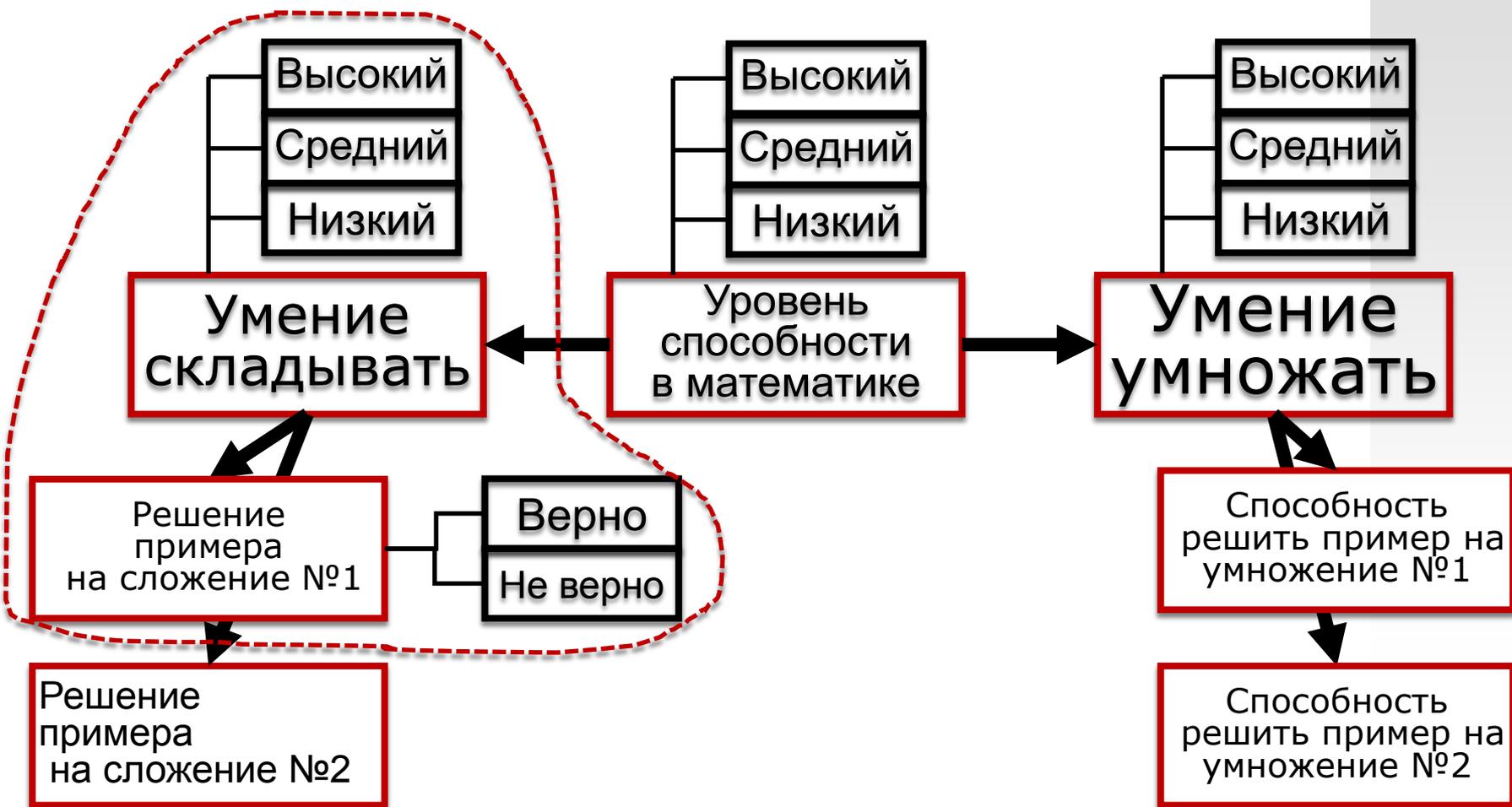
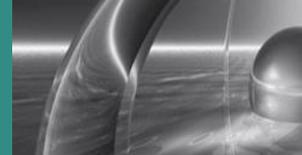
Пример сети



Пример сети



Пример сети



Условная вероятность



Метод функциональной θ

- ❖ **Проблема: при повышении количества элементов в сети, число параметров становится слишком большим**
- ❖ **Для того, чтобы обойти эту проблему разработан метод функциональной тэты (DiBello, на основе модели Samejima)**

Метод функциональной θ

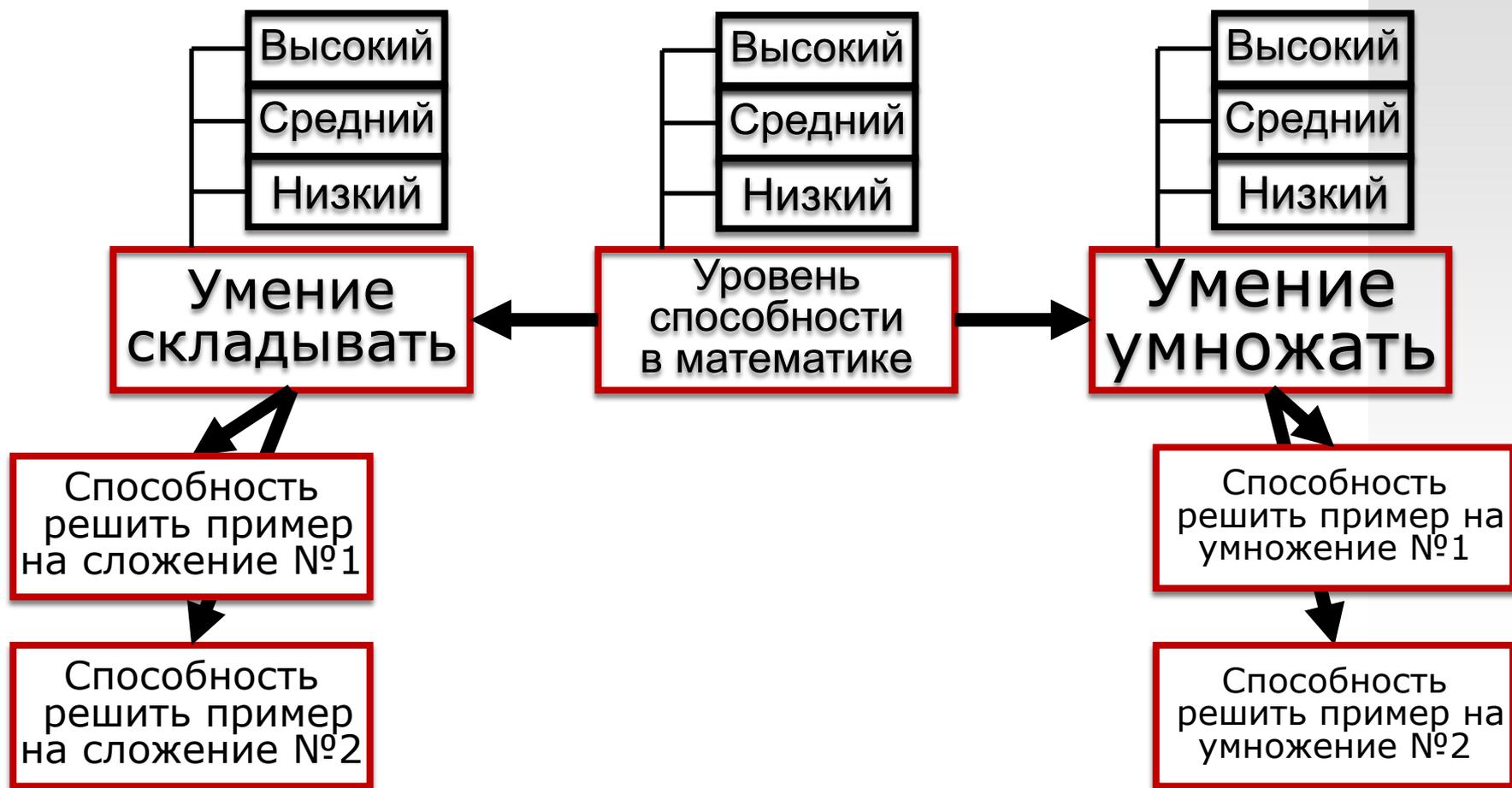
1. Для каждой наблюдаемой переменной вводим значение θ
2. Задаем количество уровней латентной переменной
3. Задаем функциональную θ для задания, $\tilde{\theta} = g(\theta_1, \dots, \theta_K)$
пользуясь функцией комбинации

$$g(\theta_1, \dots, \theta_K) = \sum_{k=1}^K \frac{\alpha_k}{\sqrt{K}} \theta_k - \beta,$$

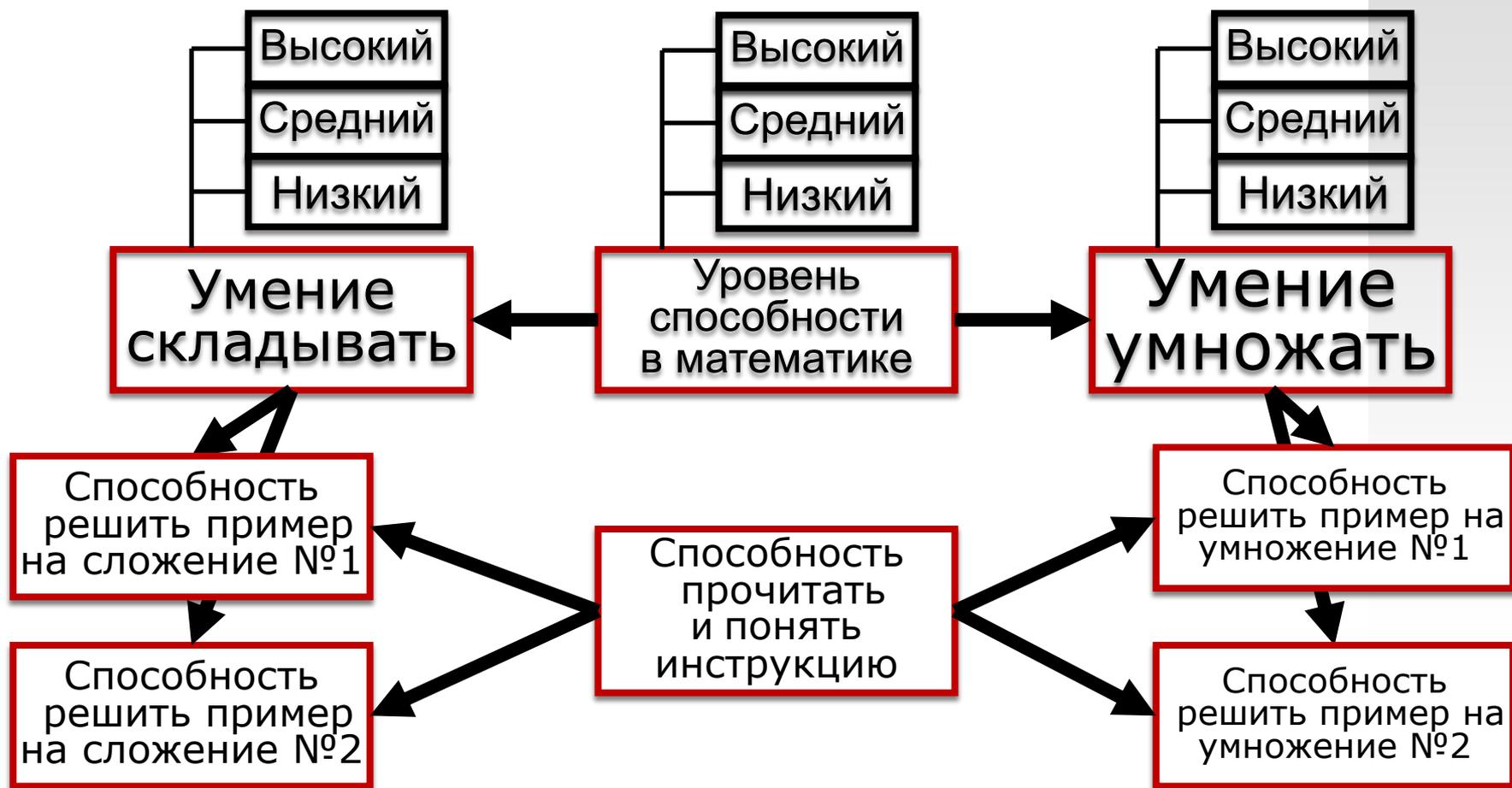
4. Используем функциональную θ для заполнения таблиц условной вероятности

$P(\text{M2_task12}|\text{MANAGE}) = \text{M2_task12} == \text{HIGH} ? 1 / (1 + \exp(-1.7 * (\text{A1}/\sqrt{1}) * \text{MANAGE} - \text{B1_2})) : \text{M2_task12} == \text{MEDIUM} ? 1 / (1 + \exp(-1.7 * (\text{A1}/\sqrt{1}) * \text{MANAGE} - \text{B1_1})) - 1 / (1 + \exp(-1.7 * (\text{A1}/\sqrt{1}) * \text{MANAGE} - \text{B1_2})) : \text{M2_task12} == \text{LOW} ? 1 - 1 / (1 + \exp(-1.7 * (\text{A1}/\sqrt{1}) * \text{MANAGE} - \text{B1_1})) : 0$

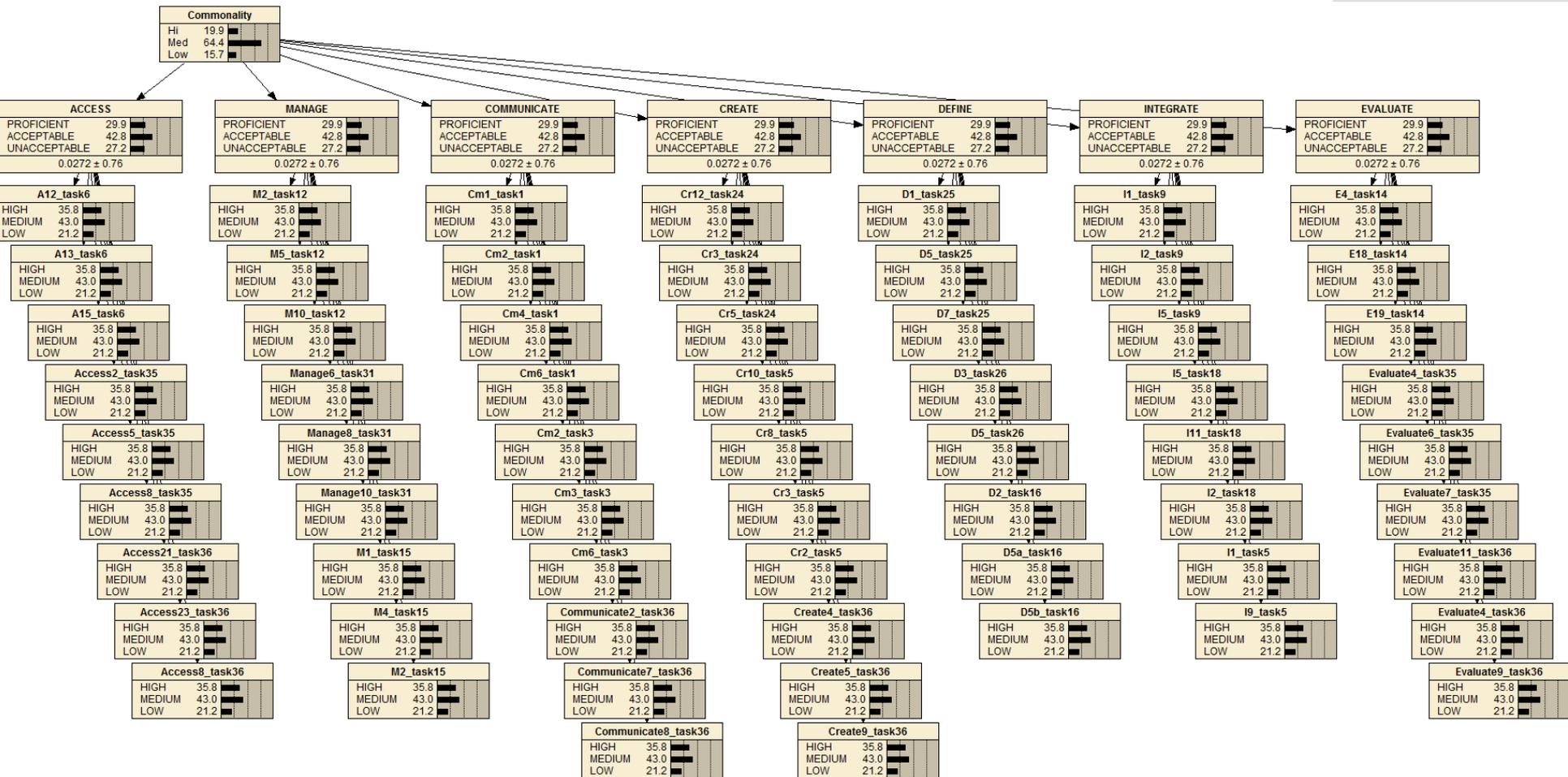
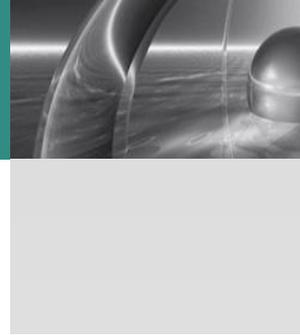
Пример сети

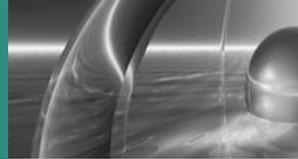


Пример сети



Сеть теста ИК-компетентности





- ❖ **Итоговый продукт сети – апостериорные распределения вероятностей всех состояний узлов, содержащих факторы ИК-компетентности**
- ❖ **Общий балл ИК-компетентности рассчитывается на основе таблицы стандартов, уже вне сети**
- ❖ **Стандарты сформированы экспертной оценкой, они позволяют регулировать относительную важность каждого фактора**



Контроль качества в байесовских сетях

Предсказание состояний узлов на симулированных данных, находящихся в полном соответствии модели

MANAGE				
Состояние узла	уверенность в предсказании			
	< 1%	< 10%	> 90%	> 99%
PROFICIENT	0,07	0,81	2,47	0,55
ACCEPTABLE	0,3	2,35	4,65	0
UNACCEPTABLE	0,25	1,39	2,21	0

В каком проценте случаев сеть неверно классифицирует?



Контроль качества измерений

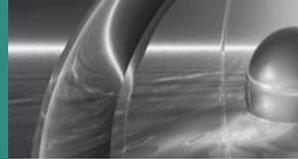
❖ Корреляционный анализ с внешними переменными:

- Самооценка ИК-компетентности
- Техническая грамотность
- Частота обращения к навыкам ИК-компетентности
- Оценки в школе

❖ Факторный анализ

❖ Анализ в IRT (с некоторыми ограничениями)

- Функционирование ответных категорий
- DIF



Плюсы в сравнении с IRT

- ❖ **Интуитивно понятное моделирование локальной зависимости**
- ❖ **Обратимость логики**
- ❖ **Моделирование взаимоотношений факторов (compensatory, conjunctive, disjunctive)**
- ❖ **Возможность работы с категориями когда это предпочтительно, без установки порогов**
- ❖ **МСМС лучше ML подходит для моделей с большим количеством факторов**



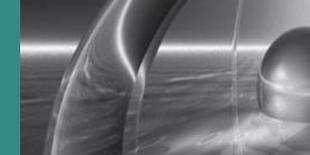
НФПК



<http://ictlit.com>

Спасибо за внимание!

Корреляционный анализ с внешними переменными



Ожидаемый коэффициент корреляции от 0.2 до 0.4

Ожидаемый коэффициент корреляции от -0.2 до 0.2

(конвергентная валидность)

(дивергентная валидность)

Собственная оценка учениками своей ИК-компетентности

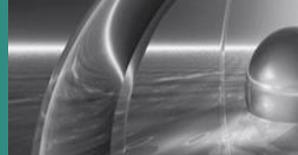
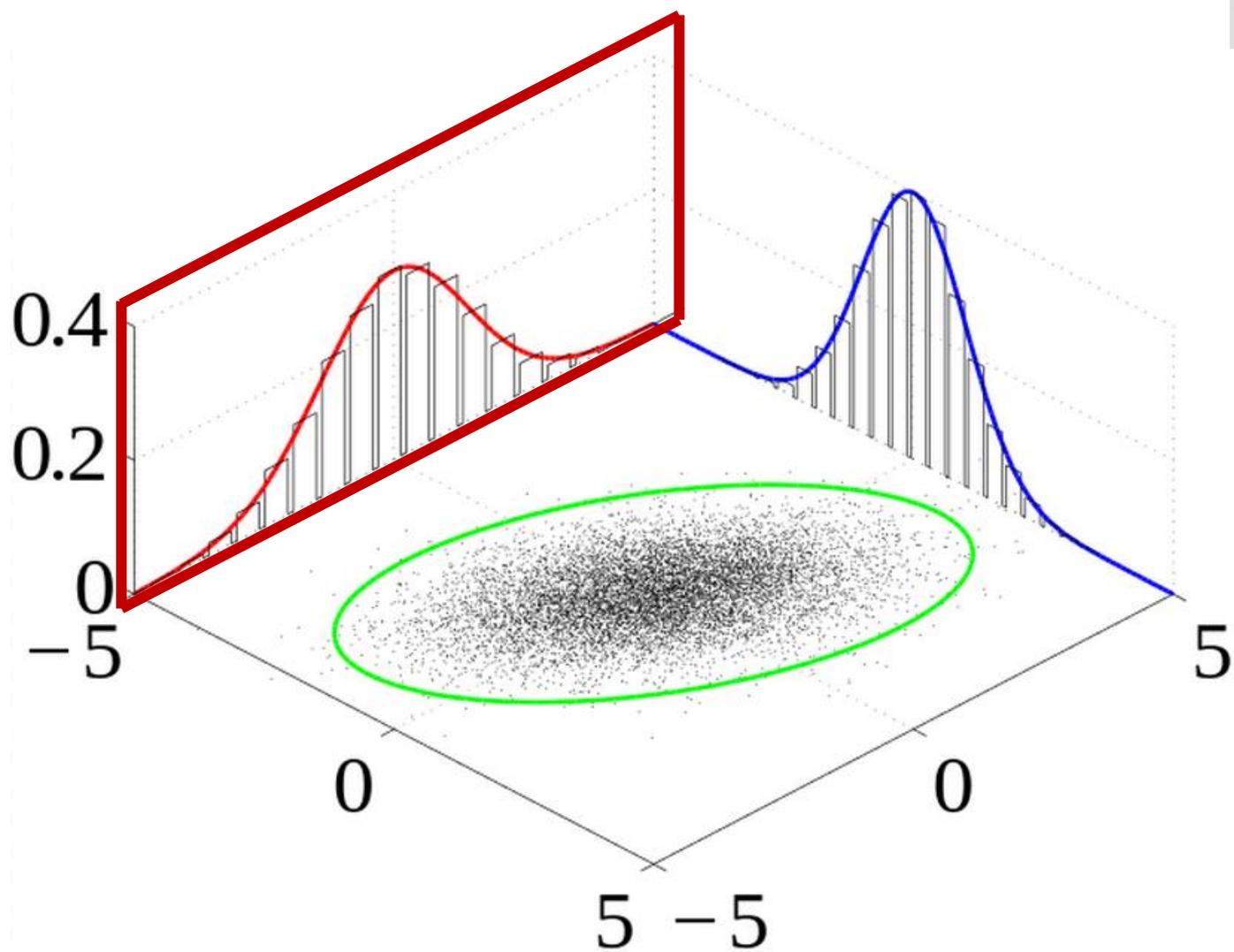
Техническая грамотность

Выборка	Кол-во заданий	Альфа	К. кор. Спирмана (стандартная ошибка)
Армения 2013	24	0.98	0.3 (0.02)*
Армения 2014	24	0.98	0.3 (0.02)*
Барнаул	23	0.97	0.28 (0.04)*
Красноярск	23	0.95	0.2 (0.05)*
Беларусь	23	0.96	0.26 (0.04)*
Москва	30	0.95	0.30 (0.11)*

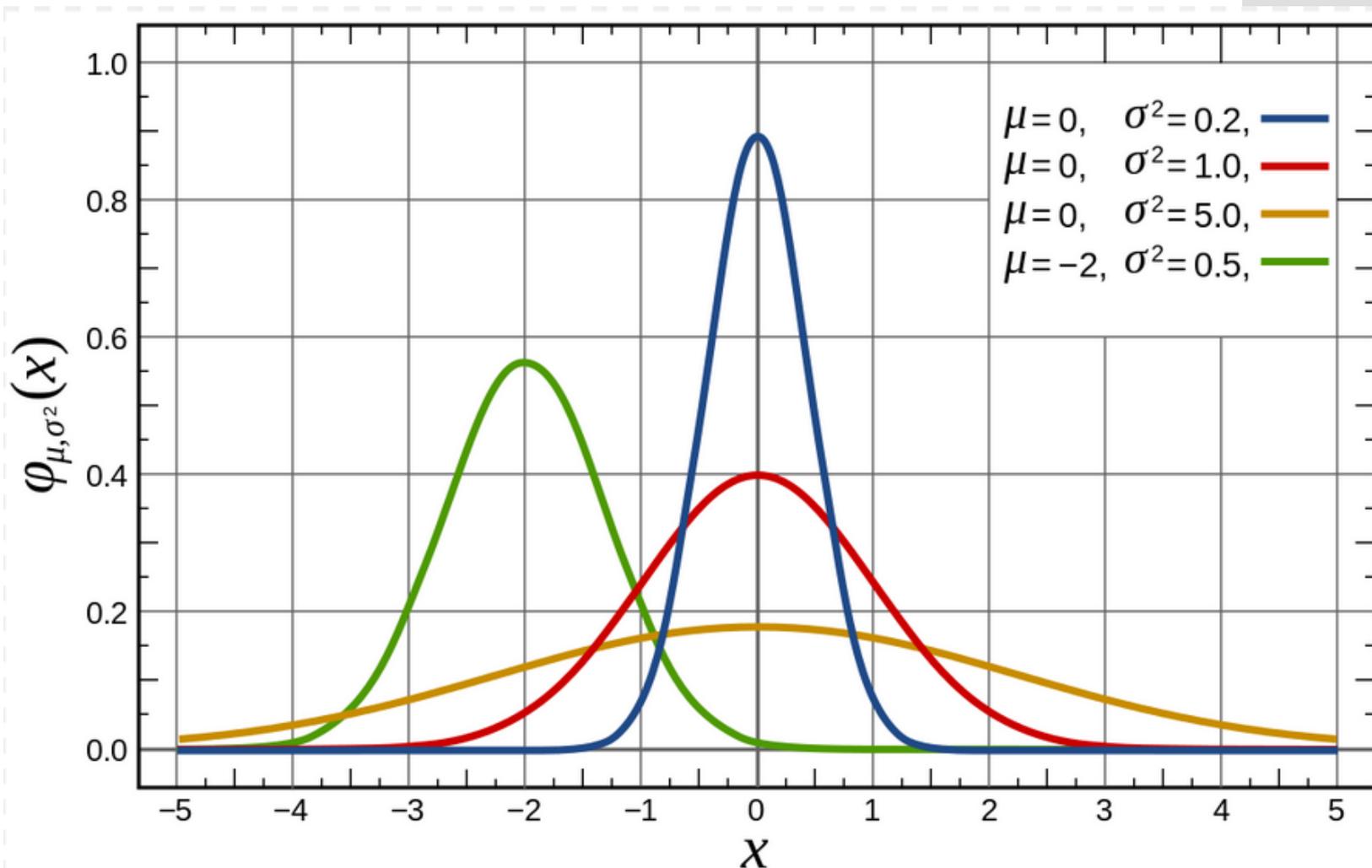
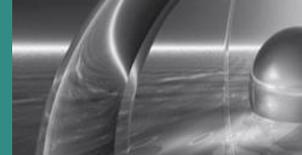
Выборка	Кол-во заданий	Альфа	К. кор. Спирмана (стандартная ошибка)
Армения 2013	12	0.93	0.18 (0.02)*
Армения 2014	12	0.93	0.18 (0.02)*
Барнаул	12	0.94	0.04 (0.04)*
Красноярск	12	0.94	0.1 (0.05)*
Беларусь	12	0.89	0.08 (0.04)*
Москва	13	0.8	0.04 (0.12)



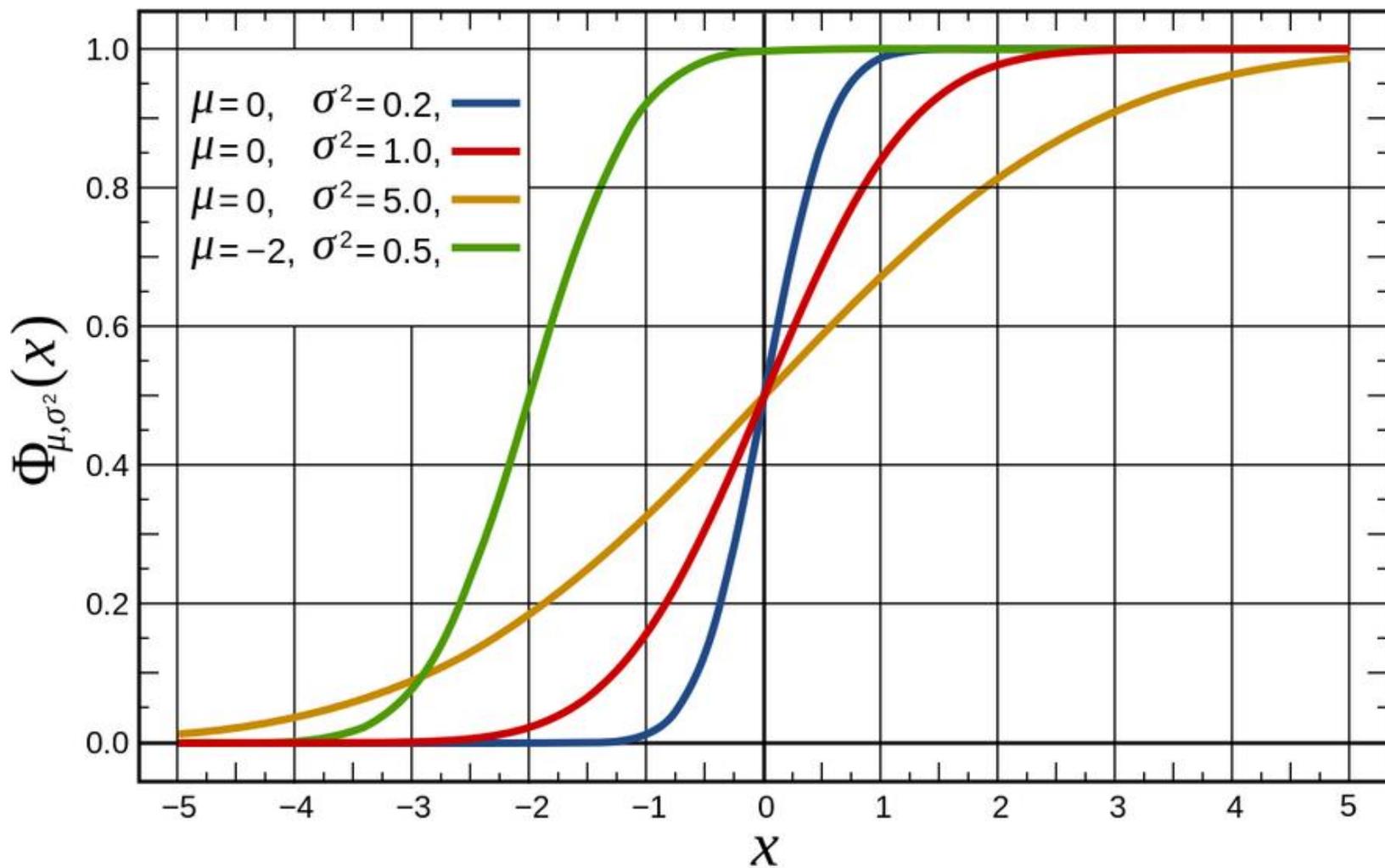
Вероятность как распределение



Гиперпараметры

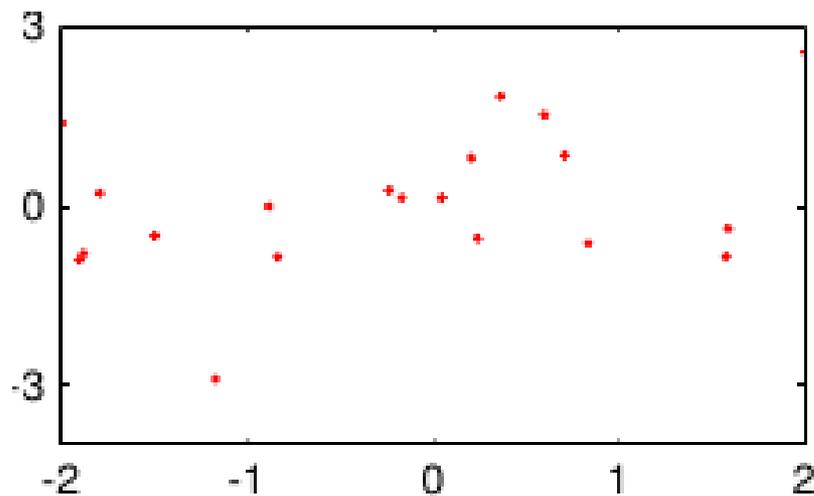


Гиперпараметры



Гиперпараметры

$$\min_{\mathbf{a} \in \mathbb{R}^{d+1}} \sum_{i=1}^d (f(x_i; \mathbf{a}) - y_i)^2 + \lambda b(\mathbf{a}).$$

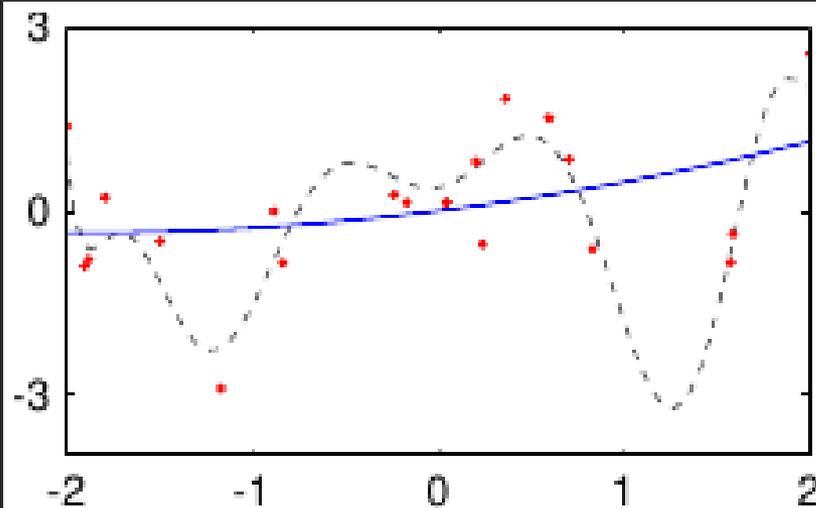


Гиперпараметры

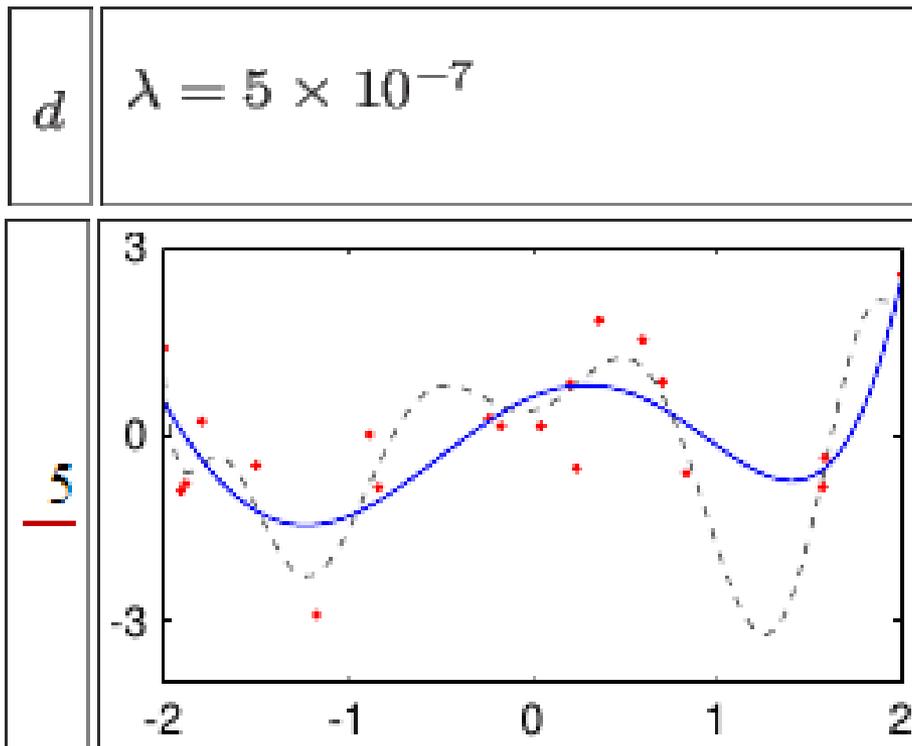
d

$$\lambda = 5 \times 10^{-7}$$

2



Гиперпараметры

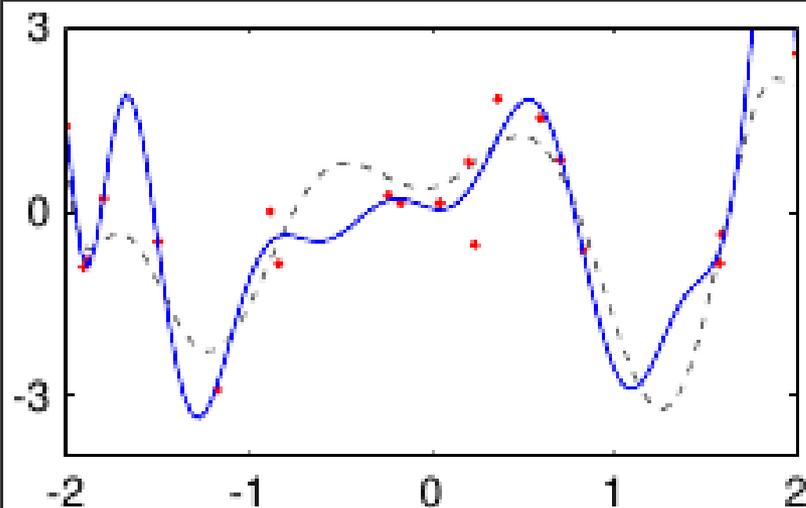


Гиперпараметры

d

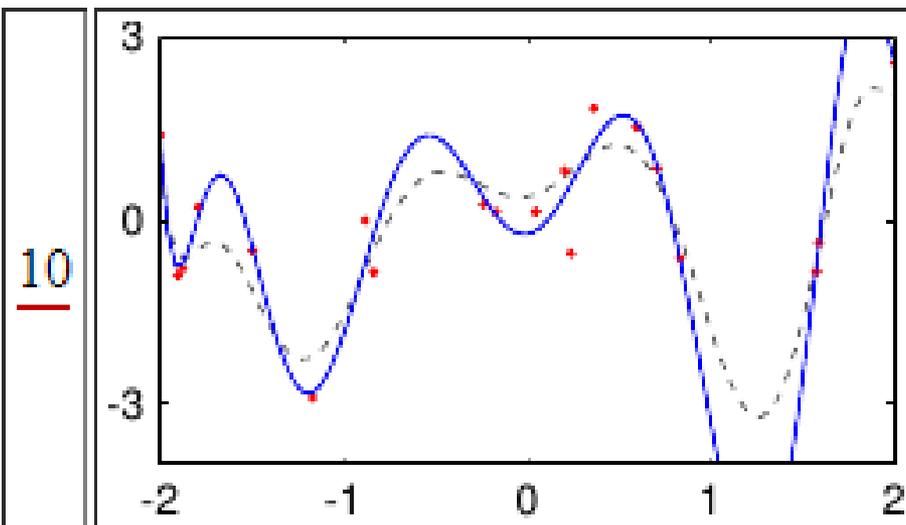
$$\lambda = 5 \times 10^{-7}$$

15



Гиперпараметры

$$d \quad \lambda = 5 \times 10^{-7}$$

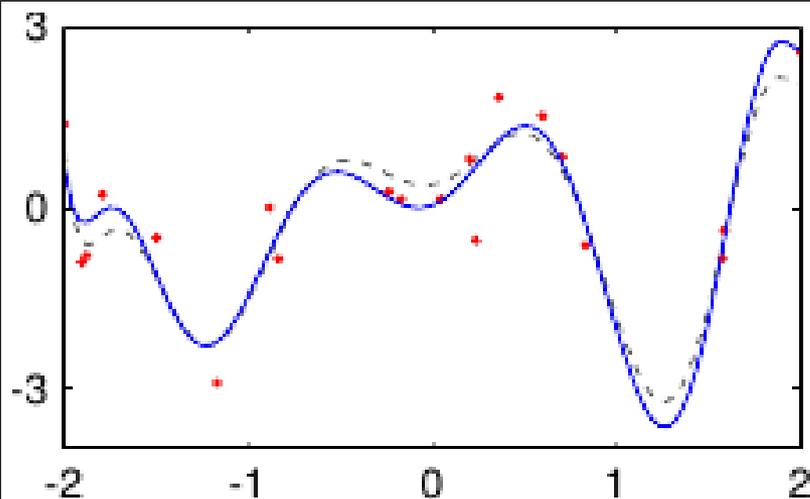


Гиперпараметры

d

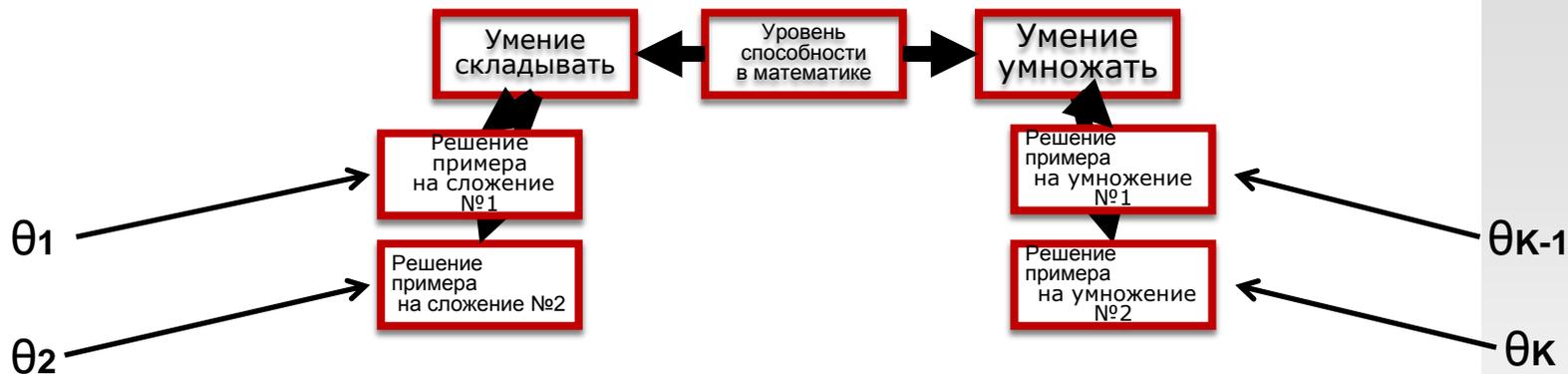
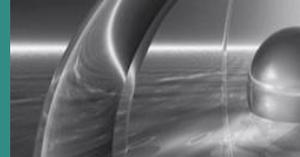
$$\lambda = 5 \times \underline{10^{-5}}$$

10



Метод функциональной

ТАТЫ



$$g(\theta_1, \dots, \theta_K) = \sum_{k=1}^K \frac{\alpha_k}{\sqrt{K}} \theta_k - \beta$$

$$Pr(X \geq x_m | \theta) = \text{logit}^{-1}(\tilde{\theta} - d_m)$$



Baye's rule states that for any two events A and C:

$$p(A | C) = \frac{p(C | A)p(A)}{P(C)} \text{ (Koski \& Noble, 2009)}$$

For this example this can be written as

$$p(X_i | Y_j) = \frac{p(Y_j | X_i)P(X_i)}{\sum_x p(X_i)p(Y_j | X_i)} \text{ where } X_i \text{ is the skill level of the student and } Y_j$$

is the outcome of the item in question.

This is often written as:

$$p(X_i | Y_j) \propto p(Y_j | X_i)P(X_i) \text{ |which can be stated as the posterior distribution}$$

(the updated probabilities of the attribute level) is proportional to the likelihood (how likely is the outcome that has been received given the prior probabilities of the attribute level) times the prior distribution (the previous belief regarding the probabilities of each attribute level).