
**СИСТЕМНЫЙ АНАЛИЗ, УПРАВЛЕНИЕ И ОБРАБОТКА ИНФОРМАЦИИ/SYSTEM ANALYSIS,
MANAGEMENT AND PROCESSING OF INFORMATION**

DOI: <https://doi.org/10.60797/IRJ.2026.170.77> EDN: VHEWZT**ДИНАМИЧЕСКАЯ МОДЕЛЬ С ЛАТЕНТНЫМИ ПАРАМЕТРАМИ ДЛЯ ОЦЕНИВАНИЯ УРОВНЯ
СФОРМИРОВАННОСТИ КОМПЕТЕНЦИЙ ПО РЕЗУЛЬТАТАМ АТТЕСТАЦИИ**

Научная статья

Родионов А.В.^{1,*}, Богорадникова Т.Г.²¹ORCID : 0000-0003-0451-2655;^{1,2} Байкальский государственный университет, Иркутск, Российская Федерация

* Корреспондирующий автор (avr-v[at]yandex.ru)

Предложена: 07.06.2026; Принята: 06.07.2026; Опубликовано: 17.08.2026

Аннотация

В компетентностном подходе компетенции выступают как обобщённый результат образования, отражающий способность обучающегося применять знания, умения и опыт деятельности в профессиональной деятельности. Процесс формирования компетенции в рамках обучения развивается постепенно, по мере прохождения обучающимся дисциплин. Поскольку уровень сформированности компетенции не может быть измерен напрямую, он оценивается косвенно, на основе результатов текущей и промежуточной аттестации. В работе этот уровень рассматривается как латентная характеристика, изменяющаяся во времени и образующая индивидуальную траекторию формирования компетенции. Цель исследования заключается в построении динамической модели, предназначенной для оценивания временной динамики латентного уровня сформированности компетенции по совокупности разнородных оценочных данных. Методологической основой служит теория ответов на задания. Параметры оцениваются методом максимума апостериорной плотности, а неопределённость оценки траектории определяется с помощью приближения Лапласа. В имитационном эксперименте с известной истинной траекторией и повторениями Монте-Карло показано, что модель восстанавливает индивидуальную траекторию формирования компетенции с малой среднеквадратической ошибкой и высокой корреляцией с истиной. Практическая значимость состоит в возможности оценивать формирование компетенций по данным текущей и промежуточных аттестаций без дополнительного тестирования.

Ключевые слова: компетенция, оценка компетенции, латентная переменная, теория ответов на задания, процесс формирования компетенции, динамическая модель оценки компетенций, оценка неопределённости.

**A DYNAMIC MODEL WITH LATENT PARAMETERS FOR EVALUATING THE LEVEL OF COMPETENCE
DEVELOPMENT BASED ON ASSESSMENT RESULTS**

Research article

Rodionov A.V.^{1,*}, Bogoradnikova T.G.²¹ORCID : 0000-0003-0451-2655;^{1,2} Baikal State University, Irkutsk, Russian Federation

* Corresponding author (avr-v[at]yandex.ru)

Suggested: 07.06.2026; Accepted: 06.07.2026; Published: 17.08.2026

Abstract

In the competence-based approach, competences are regarded as the overall outcome of education, reflecting a student's ability to apply knowledge, skills and practical experience in their professional practice. The process of competence development within the learning process unfolds gradually as students progress through their disciplines. Since the level of competence development cannot be measured directly, it is assessed indirectly, based on the results of ongoing and interim assessments. In this work, this level is regarded as a latent characteristic that changes over time and forms an individual trajectory of competence development. The aim of the research is to construct a dynamic model designed to evaluate the temporal dynamics of the latent level of competence development based on a set of heterogeneous assessment data. The methodological framework is based on item response theory. Parameters are estimated using the maximum posterior density method, and the uncertainty in the trajectory estimate is determined using the Laplace approximation. In a simulation experiment with a known true trajectory and Monte Carlo repetitions, it is shown that the model reconstructs the individual trajectory of competence development with a small mean square error and a high correlation with the true trajectory. The practical significance lies in the ability to evaluate the development of competences based on data from current and intermediary assessments without the need for additional testing.

Keywords: competence, competence assessment, latent variable, item response theory, process of competence development, dynamic model of competence assessment, uncertainty evaluation.

Введение

Переход высшей школы к компетентностной модели изменил само понимание образовательного результата — им стал не объём усвоенного материала, а компетенция, под которой понимается интегральная способность решать профессиональные задачи. Действующие стандарты ФГОС ВО определяют компетенции трёх классов

(универсальные, общепрофессиональные, профессиональные) и устанавливают для них индикаторы достижения, которые вуз закрепляет за дисциплинами и практиками образовательной программы. Принципиальная трудность оценивания состоит в том, что *уровень сформированности компетенции латентен*, иными словами, он не наблюдается непосредственно и проявляется лишь косвенно, например, через аттестацию обучающегося в определенных контрольных точках. Зачастую на практике эта латентная величина подменяется наблюдаемыми отметками, а суждение о сформированности компетенции выносится простой агрегацией — усреднением оценок по дисциплинам, формирующим компетенцию. Такой подход (расчет среднего балла или рейтинга при использовании балльно-рейтинговых систем), не учитывает динамику формирования компетенции, природу используемых измерительных шкал и неоднородность измерений [1].

Естественным аппаратом формализации подобных задач служит теория ответов на задания (Item Response Theory, IRT). Данный подход позволяет связать наблюдаемые результаты оценивания с ненаблюдаемыми характеристиками обучающегося, в частности, со сформированностью определённого образовательного результата (компетенции). В настоящее время IRT представляет целое семейство статистических моделей, используемых для анализа данных, полученных в результате тестирования, опросов или других форм оценки. Для анализа порядковых результатов оценивания, например экзаменационных оценок, используются модели градуированных и частичных оценок [2]. Для описания изменения латентного признака во времени служат продольные и динамические модели [3], [4]. В интеллектуальных обучающих системах развиваются модели отслеживания знаний, в том числе нейросетевые [5]. Однако такие модели часто обладают ограниченной интерпретируемостью, а это важный фактор для образования как сферы, в которой решения на основе искусственного интеллекта должны быть прозрачными и обоснованными, поэтому интерес представляют «объяснимые» модели и методы [6]. На этом фоне формируется представление о цифровом двойнике обучающегося как о динамической модели его образовательного состояния, которая может использоваться для мониторинга и поддержки принятия решений [7]. В ряде отечественных исследований предложены интеллектуальные модели оценивания компонентов подготовки, технологии непрерывного контроля знаний, алгоритмы управления обучением [1], а также когнитивно-диагностические модели, описывающие связь между заданиями и компетенциями.

Таким образом, существующие исследования создают теоретическую и методическую основу, однако не полностью решают задачу оценивания процесса формирования компетенции в динамике, по мере изучения дисциплин учебного плана. В большинстве случаев либо анализируется текущее состояние обучающегося, либо динамика латентного признака рассматривается без явной связи с последовательностью дисциплин, формирующих компетенцию. Кроме того, недостаточно учитывается неоднородность оценочных данных, характерная для образовательной практики, т.к. структура исходных данных об аттестации слабо формализована и обычно определяется действующими в образовательной организации нормативными документами. Например, в ФГБОУ ВО «БГУ» применяется балльно-рейтинговая система оценивания, которая предполагает накопительный принцип оценивания учебной деятельности студентов. Текущая аттестация по дисциплине включает несколько оценочных мероприятий, результаты которых суммируются в накопленный балл по стобалльной шкале. Промежуточная аттестация проводится в форме экзамена или зачёта. Экзаменационная оценка имеет порядковую пятибалльную шкалу, связанную со стобалльной: результат 91–100 баллов соответствует оценке «отлично», 76–90 баллов — «хорошо», 61–75 баллов — «удовлетворительно», менее 61 балла — «неудовлетворительно». Зачёт считается полученным при достижении результата не ниже 61 балла.

В связи с этим цель исследования заключается в построении модели, предназначенной для оценивания временной динамики латентного уровня сформированности компетенции по совокупности разнородных оценочных данных. Такая модель должна обеспечить согласованное использование результатов разных видов аттестации и позволить описывать не только итоговый уровень сформированности компетенции, но и сам процесс его изменения по мере освоения обучающимся последовательности дисциплин образовательной программы.

Методы и принципы исследования

Пусть есть компетенция, для которой образовательная программа задаёт упорядоченный список из D дисциплин. Обозначим через $i = 1, \dots, N$ номер обучающегося, а через $d = 1, \dots, D$ номер дисциплины в этой последовательности. Учебное время обозначим через $\tau \in [0, \tau_{\max}]$, где $\tau_{\max}$ соответствует завершению рассматриваемого фрагмента образовательной программы. По каждой дисциплине фиксируются результаты текущей и промежуточной аттестации. Текущая аттестация по дисциплине складывается из результатов отдельных контрольных мероприятий, определенных в фонде оценочных средств. Результаты этой аттестации представляют собой вектор наблюдений, сумма баллов — 100. Промежуточная аттестация представляет собой итоговую оценку освоения дисциплины в пятибалльной или двухбалльной шкале. Задача состоит в том, чтобы на основе доступных аттестационных данных оценить индивидуальную траекторию латентного уровня сформированности компетенции $\theta_i(\tau)$ изменяющегося по мере освоения обучающимся дисциплин образовательной программы.

Предлагаемая в работе модель описывает формирование компетенции как динамический латентный процесс. Латентный уровень сформированности компетенции студента в момент учебного времени рассматривается как сумма сглаженного индивидуального тренда и краткосрочной случайной флуктуации:

$$\theta_i(\tau) = \mu_i(\tau) + u_i(\tau), \quad \tau \in [0, \tau_{\max}], \quad (1)$$

где $\mu_i(\tau)$ описывает основную траекторию формирования компетенции, а $u_i(\tau)$ отражает случайные отклонения от неё. Для задания тренда используется логистическая функция вида

$$\Lambda(x) = \frac{1}{1 + \exp(-x)}, \quad x \in \mathbb{R}. \quad (2)$$

Чтобы параметры модели имели содержательную интерпретацию начального и конечного уровней на рассматриваемом интервале обучения, используется нормированная логистическая функция

$$G(\tau) = \frac{\Lambda(\gamma(\tau - \tau^*)) - \Lambda(-\gamma\tau^*)}{\Lambda(\gamma(\tau_{\max} - \tau^*)) - \Lambda(-\gamma\tau^*)}. \quad (3)$$

Тогда индивидуальный тренд задаётся выражением

$$\mu_i(\tau) = \theta_i^0 + \Delta_i G(\tau), \quad \Delta_i = \theta_i^\infty - \theta_i^0. \quad (4)$$

где θ_i^0 соответствует начальному уровню сформированности компетенции, θ_i^∞ уровню к завершению рассматриваемого периода обучения, а Δ_i характеризует общий прирост латентного уровня. Параметр $\gamma > 0$ определяет скорость формирования компетенции, а $\tau^* \in (0, \tau_{\max})$ задаёт момент наиболее интенсивного изменения тренда. Краткосрочные отклонения от тренда описываются процессом Орнштейна-Уленбека:

$$du_i(\tau) = -\kappa u_i(\tau)d\tau + \sigma_u dW_i(\tau), \quad (5)$$

где $\kappa > 0$ параметр возврата к среднему уровню, $\sigma_u > 0$ интенсивность случайных флуктуаций, а $W_i(\tau)$ стандартный винеровский процесс. Процесс Орнштейна-Уленбека рассматривается в стационарном режиме и выбран как непрерывный аналог авторегрессии первого порядка, описывающий затухающие во времени колебания результативности с возвратом к среднему уровню. Латентным аргументом измерительной модели в каждый момент наблюдения служит само значение траектории — сумма гладкого тренда и краткосрочной флуктуации

$$\eta_i(\tau) \equiv \theta_i(\tau) = \mu_i(\tau) + u_i(\tau) \quad (6)$$

где первое слагаемое — гладкий тренд, а второе — краткосрочная флуктуация Орнштейна—Уленбека. Результаты аттестации описываются в рамках единой измерительной модели. Каждый результат контрольного мероприятия текущей аттестации по стобалльной шкале рассматривается как непрерывная наблюдаемая величина, привязанная к своему моменту учебного времени:

$$T_{d,j} | \eta_i(\tau_{d,j}) \sim \mathcal{N}(m_{d,j}^C, \sigma_C^2), \quad (7)$$

где $m_{d,j}^C = q_{d,j} \Lambda(a^C(\eta_i(\tau_{d,j}) - b^C))$, $q_{d,j}$ — максимальный балл за j -е контрольное мероприятие дисциплины d , $\sum_{j \in J_d} q_{d,j} = 100$, параметр a^C определяет чувствительность текущего балла к изменению латентного уровня b^C задаёт положение шкалы, σ_T^2 характеризует дисперсию ошибки измерения.

Экзаменационная оценка Y_{id}^E описывается как порядковая категориальная переменная с четырьмя уровнями. Категории кодируются числами 0,1,2,3, где 0 соответствует оценке «неудовлетворительно», а 3 оценке «отлично». Для экзамена используется модель градуированных ответов:

$$PY_{id}^E \geq k | \eta_{id} = \frac{\exp(a^E(\eta_{id} - b_k^E))}{1 + \exp(a^E(\eta_{id} - b_k^E))}. \quad (8)$$

Результат зачёта $Y_{id}^Z \in 0, 1$ моделируется двухпараметрической моделью:

$$PY_{id}^Z = 1 | \eta_{id} = \frac{\exp(a^Z(\eta_{id} - b^Z))}{1 + \exp(a^Z(\eta_{id} - b^Z))}, \quad (9)$$

где a^Z характеризует дискриминирующую способность зачётного измерения, а b^Z — порог латентного уровня, соответствующий получению зачёта.

Для оценки параметров модели воспользуемся методом максимума апостериорной плотности (Maximum a Posteriori, MAP). Введём множества $\mathcal{E}_i$ и $\mathcal{Z}_i$. Множество $\mathcal{E}_i$ содержит дисциплины с наблюдаемой экзаменационной оценкой, а $\mathcal{Z}_i$ дисциплины, завершающиеся зачётом. Индивидуальный вектор оцениваемых параметров запишем в виде

$$\varphi_i = (\theta_i^0, \Delta_i, u_i(\tau_1), \dots, u_i(\tau_n)). \quad (10)$$

Логарифм апостериорной плотности для студента i задаётся выражением

$$\ell(\varphi_i) = \sum_{d=1}^D \sum_{j \in J_d} \log f_N(T_{d,j}, m_{d,j}^C, \sigma_C^2) + \sum_{d \in \mathcal{E}_i} \log P(Y_{id}^E = y_{id}^E | \eta_{id}) + \sum_{d \in \mathcal{Z}_i} \log P(Y_{id}^Z = z_{id}^Z | \eta_{id}) + \log p(\varphi_i). \quad (11)$$

В качестве априорных распределений используются нормальные распределения для начального уровня и прироста:

$$\theta_i^0 \sim N(m_0, s_0^2), \Delta_i \sim N(m_\Delta, s_\Delta^2), \quad (12)$$

а для состояний процесса Орнштейна-Уленбека в дискретных моментах наблюдения задаётся нормальное распределение

$$u_i \sim \mathcal{N}(0, K_i), \quad [K_i]_{kl} = \frac{\sigma_u^2}{2\kappa} e^{-\kappa|\tau_k - \tau_l|}. \quad (13)$$

Тогда оценка индивидуальных параметров находится:

$$\hat{\varphi}_i = \arg \max_{\varphi_i} \ell_i(\varphi_i). \quad (14)$$

Неопределённость оценки определяется с использованием приближения Лапласа. Ковариационная матрица оценки задаётся выражением

$$\hat{\Sigma}_i = [-\nabla^2 \ell_i(\hat{\varphi}_i)]^{-1}. \quad (15)$$

Дисперсия оценки тренда в произвольный момент учебного времени вычисляется по дельта-методу:

$$\text{Var} \{ \hat{\mu}_i(\tau) \} \approx \nabla_{\varphi} \mu_i(\tau)^T \hat{\Sigma}_i \nabla_{\varphi} \mu_i(\tau). \quad (16)$$

Для выбранной параметризации градиент имеет вид

$$\nabla_{\varphi} \mu_i(\tau) = (1, G(\tau), 0, \dots, 0)^T. \quad (17)$$

На основе этой дисперсии строится 95-процентный апостериорный интервал для индивидуального тренда:

$$\hat{\mu}_i(\tau) \pm 1,96 \sqrt{\text{Var} \{ \hat{\mu}_i(\tau) \}}. \quad (18)$$

Для исследования предлагаемой в работе модели воспользуемся имитационным моделированием. Имитационная модель должна одновременно порождать эталонную траекторию каждого студента и наблюдаемые результаты аттестации. Чтобы отделить закономерные различия методов от случайности конкретной выборки, эксперимент независимо повторяется множество раз (Монте-Карло), и все показатели усредняются по повторениям.

Истинное значение трендовой составляющей латентного уровня обозначим μ_{it} , а его оценку $\hat{\mu}_{it}$. Среднеквадратическая ошибка $RMSE = \sqrt{\frac{1}{N \cdot T} \sum_{i,t} (\hat{\mu}_{it} - \mu_{it})^2}$ характеризует среднюю величину отклонения оценки от истины и сильнее штрафует крупные ошибки, поэтому служит основным показателем точности. Смещение $\delta = \frac{1}{N \cdot T} \sum_{i,t} (\hat{\mu}_{it} - \mu_{it})$ выявляет устойчивую недооценку или переоценку уровня, например из-за «потолка» стобальных шкал. Коэффициент корреляции $r = \frac{\text{cov}(\hat{\mu}, \mu)}{\sigma_{\hat{\mu}} \sigma_{\mu}}$ измеряет согласие формы оценённой и истинной траекторий и инвариантен к сдвигу и масштабу, поэтому показывает, воспроизводится ли сам ход формирования компетенции, даже если оценки смещены. Эмпирическое покрытие $C_{0,95} = \frac{1}{N \cdot T} \sum_{i,t} [\mu_{it} \in \hat{I}_{it}]$ представляет долю истинных значений, попавших в 95-процентный апостериорный интервал $\hat{I}_{it}$.

Адекватность оценённой модели проверяется по остаткам наблюдаемых результатов аттестации относительно их модельных ожиданий. Для каждого наблюдаемого результата вычисляется

$$z_{id} = \frac{x_{id} - M[X_{id}]}{\sqrt{D[X_{id}]}}, \quad (19)$$

На основе стандартизованных остатков для каждого обучающегося (и аналогично для каждой дисциплины) рассчитываются невзвешенная (Outfit) и взвешенная по дисперсии (Infit) статистики согласия [8]:

$$U_i = \frac{1}{|J_i|} \sum_{d \in J_i} z_{id}^2, \quad (20)$$

$$V_i = \frac{\sum_{d \in J_i} (x_{id} - M[X_{id}])^2}{\sum_{d \in J_i} D[X_{id}]}. \quad (21)$$

Значения статистик, близкие к единице, свидетельствуют о согласии данных с моделью, а значения, заметно превышающие единицу, указывают на нетипичные результаты оценивания. Долю объяснённой моделью изменчивости оценок характеризует коэффициент детерминации [9]:

$$R^2 = 1 - \frac{\sum_{i,d} (x_{id} - M[X_{id}])^2}{\sum_{i,d} (x_{id} - \bar{x})^2}. \quad (22)$$

Согласие отдельных измерений с моделью оценивается статистиками типа S–X² Орландо-Тиссена и Q1 Йена [10]. Обучающиеся разбиваются на группы по оценённому латентному уровню (статистика Q1) либо по наблюдаемому суммарному баллу (статистика S–X²), и в каждой группе g наблюдаемые частоты категорий сопоставляются с ожидаемыми по модели:

$$S - X_j^2 = \sum_{g=1}^G \sum_c \frac{(o_{gc} - E_{gc})^2}{E_{gc}}, \quad (23)$$

где O_{gc} — число обучающихся группы g , получивших категорию c по измерению j , E_{gc} — соответствующее модельное ожидание, G — число групп. Полученная статистика имеет приближённое распределение хи-квадрат, и значимое отклонение указывает на несоответствие измерительной модели для данного мероприятия. Предположение о локальной независимости измерений проверяется индексом Q3 Йена [10]:

$$Q_3(j, k) = \frac{\sum_i (z_{ij} - \bar{z}_j)(z_{ik} - \bar{z}_k)}{\sqrt{\sum_i (z_{ij} - \bar{z}_j)^2 \sum_i (z_{ik} - \bar{z}_k)^2}}. \quad (24)$$

При выполнении условной независимости значения Q3 близки к небольшой отрицательной величине, тогда как заметные положительные значения свидетельствуют о локальной зависимости.

Для сравнения вариантов модели, в частности для проверки необходимости компоненты краткосрочной флуктуации, используются информационные критерии [11]. Критерий Акаике и байесовский информационный критерий вычисляются по логарифму правдоподобия при оценённых параметрах: $AIC = -2 \ln L + 2k$ и $BIC = -2 \ln L + k \ln N_{obs}$, где L — правдоподобие, N_{obs} — число наблюдений, k — эффективное число параметров. Широко применимый информационный критерий оценивает предсказательную способность по апостериорному распределению:

$$\begin{aligned} lppd &= \sum_{i,j} \ln \left(\frac{1}{S} \sum_{s=1}^S p(y_{ij} | \theta_i^{(s)}) \right) \\ p_{WAIC} &= \sum_{i,j} \text{Var}_s \left(\ln p(y_{ij} | \theta_i^{(s)}) \right) \\ WAIC &= -2 (lppd - p_{WAIC}) \end{aligned} \quad (25)$$

Поскольку латентные параметры регуляризованы априорными распределениями, эффективное число параметров определяется следом матрицы влияния $k = n - \text{tr}(C\Lambda)$, где C — апостериорная ковариационная матрица оценки, Λ — матрица точности априорного распределения; оно оказывается меньше номинального числа латентных величин. Меньшим значениям критериев отвечает предпочтительный вариант модели.

Основные результаты

Для проверки внутренней корректности модели проведён имитационный эксперимент. Имитационное моделирование, оценивание параметров и расчёт метрик выполнены программным модулем, который разработан авторами на языке Python. Для численных вычислений использованы библиотеки NumPy и SciPy. Операции линейной алгебры векторизованы, а псевдослучайные числа порождаются генератором PCG64 с фиксированными начальными значениями, что обеспечивает воспроизводимость. Краткосрочная флуктуация моделируется процессом Орнштейна-Уленбека и оценивается как состояние: латентный вектор включает её значения в моменты наблюдений с марковским (точно дискретизированным) априорным распределением. Логистическая функция вычисляется функцией `scipy.special.expit`. Максимум апостериорной плотности для каждого студента отыскивается квазиньютоновским методом L-BFGS-B из `scipy.optimize.minimize`. Логарифм плотности задаётся в замкнутой форме, а его градиент вычисляется численно. Матрица вторых производных в точке максимума находится конечно-разностным методом с учётом её разрежённой (стрелочно-трёхдиагональной) структуры, а ковариация параметров и апостериорные интервалы траектории получаются её обращением средствами NumPy. Визуализация выполнена средствами библиотеки Matplotlib.

Компетенция формировалась последовательностью из десяти дисциплин, по две в каждом из пяти семестров. Семь дисциплин завершались экзаменом, а три — зачётом. По каждой дисциплине генерировались вектор результатов контрольных мероприятий текущей аттестации (по две контрольные точки) и итоговая оценка промежуточной аттестации согласно измерительной модели. Параметры эксперимента приведены в табл. 1 и подобраны так, чтобы воспроизвести типичные распределения реальных оценок. Каждый сценарий воспроизводился в нескольких независимых повторениях (базовый — 200, сценарии с нарушением предпосылок — по 100), в каждом из которых моделировалась когорта из 50 обучающихся. Для каждой метрики приводятся её среднее значение и среднее квадратическое отклонение по повторениям. Пример структуры исходных данных и измерительной модели представлены на рис. 1.

Таблица 1 - Параметры базового имитационного эксперимента

DOI: <https://doi.org/10.60797/IRJ.2026.170.77.1>

Параметр	Значение
Число студентов N	50
Число дисциплин D (экзамен / зачёт)	10 (7 / 3)
Семестров (дисциплин в семестр)	5 (2)
Шкала текущей аттестации	0–100 (непрерывная)
Категории экзамена (установленные диапазоны)	неуд / удовл / хор / отл
Дискриминирующие способности: $a^C / a^E / a^Z$	1,2 / 1,4 / 1,4
Пороги экзамена (GRM): $b^E_1 / b^E_2 / b^E_3$	–0,5 / 0,5 / 1,4
Параметры положения: b^C / b^Z	–0,3 / –1,0
Начальный уровень θ^0 (среднее / СКО)	–0,5 / 0,5
Прирост $\theta^\infty - \theta^0$ (среднее / СКО)	2,0 / 0,5
Параметры тренда: γ / τ^*	2,5 / 1,2
Параметры флуктуации: κ / σ_u	1,5 / 0,45
СКО контрольной точки σ_C	9,0
Контрольных точек текущей аттестации на	2

Параметр	Значение
дисциплину	

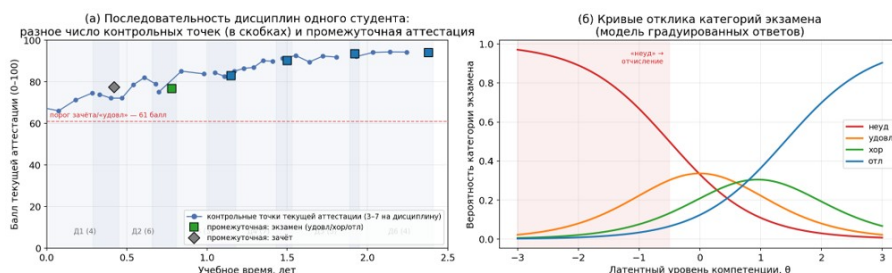


Рисунок 1 - Структура данных и измерительная модель

DOI: <https://doi.org/10.60797/IRJ.2026.170.77.2>

Примечание: (а) показана последовательность дисциплин одного студента с вектором контрольных точек текущей аттестации и промежуточной аттестацией в виде категории экзамена или зачёта; (б) приведены кривые отклика категорий экзамена в модели градуированных ответов, то есть вероятности категорий «неуд», «удовл», «хор» и «отл» как функции латентного уровня компетенции

Предложенная модель оценивает индивидуальную траекторию формирования компетенции вместе с доверительным интервалом (рис. 2). Оценённая траектория близка к истинной на всём интервале обучения, а апостериорная полоса расширяется там, где данных меньше, отражая рост неопределённости. В базовом сценарии модель оценила траектории со средней квадратической ошибкой $0,227 \pm 0,030$ и корреляцией 0,976, а эмпирическое покрытие 95-процентных апостериорных интервалов составило 0,932. Это означает, что фактическое покрытие близко к номинальному уровню и приближение Лапласа адекватно оценивает неопределённость.

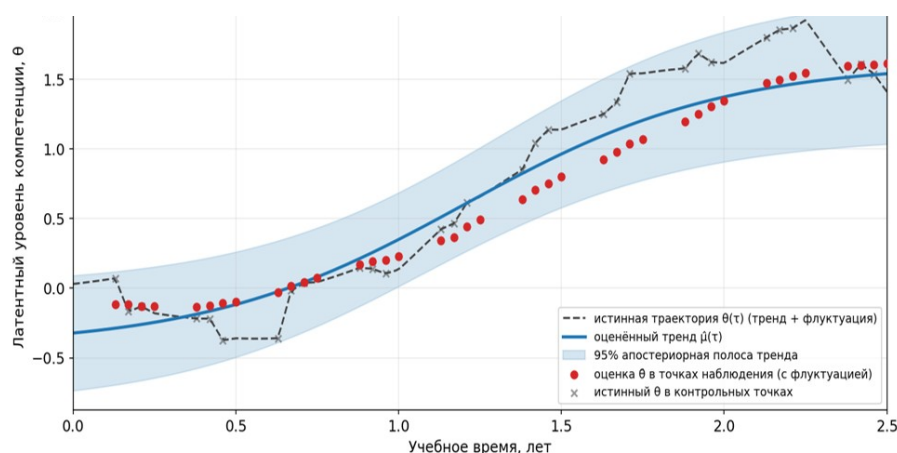


Рисунок 2 - Показаны истинная и оценённая траектории формирования компетенции с 95-процентной апостериорной полосой

DOI: <https://doi.org/10.60797/IRJ.2026.170.77.3>

Для проверки устойчивости предложенной модели рассмотрены отклонения от модельных предпосылок (табл. 2).

Таблица 2 - Устойчивость предложенной модели к нарушению предпосылок

DOI: <https://doi.org/10.60797/IRJ.2026.170.77.4>

Сценарий	RMSE	Корр.	Смещение
Базовый сценарий	$0,227 \pm 0,030$	0,976	$-0,061 \pm 0,026$
Повышенная дисперсия контрольной точки ($\sigma_C = 14$)	$0,274 \pm 0,022$	0,962	$-0,060 \pm 0,037$
Неверно заданные априорные	$0,229 \pm 0,023$	0,976	$-0,056 \pm 0,026$

Сценарий	RMSE	Корр.	Смещение
распределения			

При повышении дисперсии контрольной точки (σ_C с 9 до 14) и при заведомо неверно заданных (смещённые средние начального уровня и прироста, удвоенные стандартные отклонения) средняя квадратическая ошибка возрастает умеренно, с 0,227 до 0,274 и 0,229 соответственно. Это показывает, что корректность модели не сводится к совпадению априорных распределений с механизмом порождения данных и сохраняется при их рассогласовании и при росте шума измерений.

Применимость предложенной модели проиллюстрируем на примере общепрофессиональной компетенции образовательной программы бакалавриата по направлению 09.03.03 «Прикладная информатика» ФГБОУ ВО «БГУ». Рассмотрим компетенцию ОПК-1 — способность применять естественнонаучные и общетехнические знания, методы математического анализа и моделирования, теоретического и экспериментального исследования в профессиональной деятельности. Перечень дисциплин, формирующих компетенцию, приведен в таблице 3.

Таблица 3 - Дисциплины, формирующие компетенцию ОПК-1

DOI: <https://doi.org/10.60797/IRJ.2026.170.77.5>

Дисциплина	Семестр	Промежуточная аттестация
Линейная алгебра	1	зачёт
Математический анализ	2	экзамен
Численные методы	3	экзамен
Теория вероятностей и математическая статистика	4	экзамен
Исследование операций	5	экзамен

Для анализа сформирован датасет из 43 обучающихся по пяти дисциплинам. По каждой дисциплине строился вектор результатов текущей аттестации и одна оценка промежуточной аттестации, всего 1290 наблюдений. Оценённые траектории латентного уровня сформированности компетенции для трёх типичных профилей: со стабильно высокими оценками («отлично»); с колебаниями между «отлично» и «хорошо»; со стабильными «хорошо» приведены на рис. 3. Статистики согласия Outfit и Infit по всем пяти дисциплинам находятся в пределах допустимого диапазона 0,5–1,5 (Outfit от 0,58 до 0,88, Infit от 0,60 до 0,94), рис. 4.

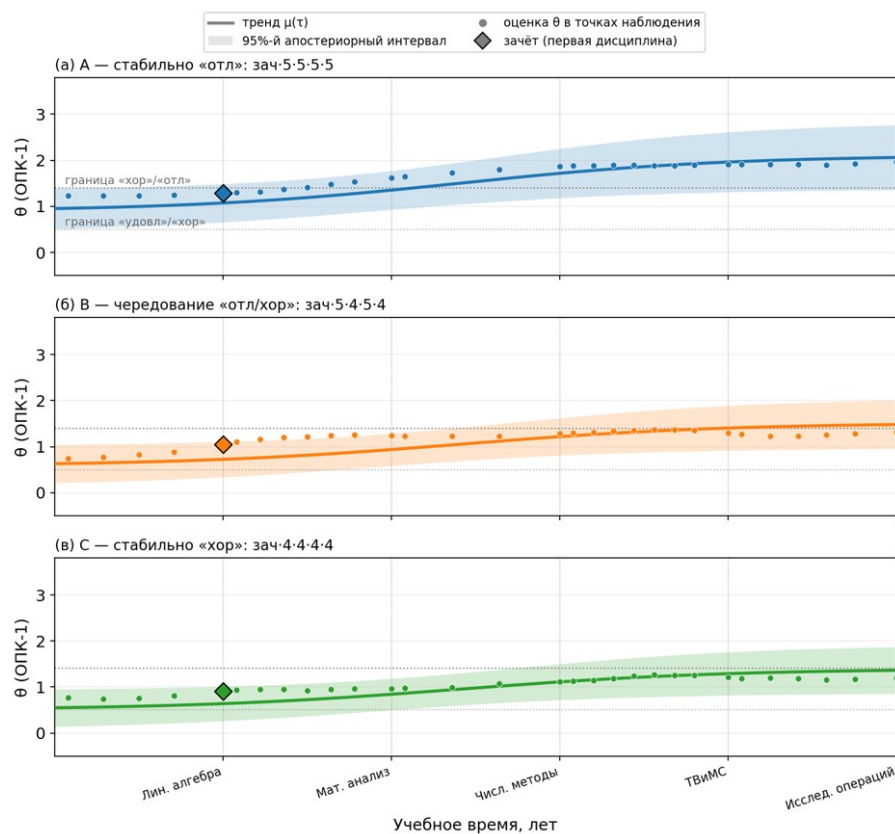


Рисунок 3 - Оценённые траектории латентного уровня сформированности ОПК-1 для трёх типичных профилей оценок

DOI: <https://doi.org/10.60797/IRJ.2026.170.77.6>

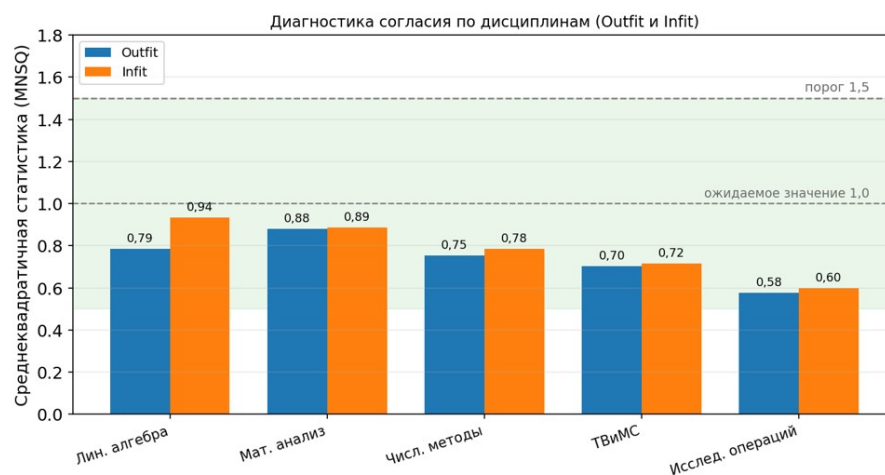


Рисунок 4 - Статистики согласия Outfit и Infit по дисциплинам последовательности для примера ОПК-1

DOI: <https://doi.org/10.60797/IRJ.2026.170.77.7>

Работоспособность модели дополнительно проверена тремя группами диагностик (табл. 4).

Таблица 4 - Диагностика работоспособности модели

DOI: <https://doi.org/10.60797/IRJ.2026.170.77.8>

Критерий	Полученные значения	Допустимые значения	Интерпретация
S-X ² (Орландо—Тиссен), Q1 (Йена)	p от 0,14 до 0,91	p > 0,05	Категориальные частоты согласуются с моделью,

Критерий	Полученные значения	Допустимые значения	Интерпретация
			систематического несоответствия нет
Q3 (Йена) — локальная зависимость	Среднее по парам 0,10; макс. по модулю 0,33 (после центрирования 0,25); «текущая–промежуточная» от –0,30 до 0,24	$ Q3 \leq 0,2$	В среднем выраженной локальной зависимости не выявлено, однако отдельные пары измерений демонстрируют значения Q3 выше ориентировочного порога и требуют дополнительной проверки
AIC, WAIC (полная и упрощённая модель)	$\Delta AIC \approx 35$; $\Delta WAIC \approx 37$ в пользу полной модели	$\Delta > 0$ в пользу модели ($\Delta > 10$ — сильное предпочтение)	Критерии AIC и WAIC свидетельствуют в пользу полной модели с краткосрочной динамической компонентой, тогда как BIC, сильнее штрафующий сложность, отдаёт предпочтение упрощённой модели
BIC (полная и упрощённая модель)	В пользу упрощённой модели	$\Delta > 0$ в пользу модели	

Обсуждение

Полученные результаты подтверждают целесообразность рассмотрения сформированности компетенции как динамической латентной характеристики, изменяющейся по мере освоения обучающимися последовательности дисциплин образовательной программы. Предложенная модель позволяет перейти от простой агрегации наблюдаемых оценок к реконструкции индивидуальной траектории формирования компетенции, учитывающей временную структуру обучения и различную измерительную природу аттестационных данных. Существенным преимуществом является согласованное включение в единую вероятностную схему накопленных баллов текущей аттестации, порядковых экзаменационных оценок и бинарных результатов зачёта.

Имитационный эксперимент показал, что модель восстанавливает латентную траекторию с малой среднеквадратической ошибкой, высокой корреляцией с истинными значениями и близким к номинальному покрытию апостериорных интервалов. Устойчивость результатов при увеличении дисперсии контрольных точек и при рассогласовании априорных распределений свидетельствует о том, что работоспособность модели не сводится к искусственному совпадению условий моделирования и оценивания. Используемые в работе статистические критерии Outfit, Infit, $S-X^2$, Q1 и Q3 показывают отсутствие выраженного систематического несоответствия модели данным в рассмотренном примере.

Заключение

Предложена динамическая модель оценивания латентной траектории формирования одной компетенции по результатам текущей и промежуточной аттестации. Латентный уровень сформированности описан непрерывной траекторией, связанной с последовательностью формирующих дисциплин. Разнородные шкалы оценивания, принятые в вузе и включающие накопленный балл текущей аттестации, категориальную оценку экзамена и бинарный результат зачёта, согласованы единой измерительной моделью, а оценка траектории дополнена количественной мерой неопределённости.

Практическая значимость модели определяется возможностью применения уже имеющихся в университете аттестационных процедур без обязательного введения дополнительных процедур тестирования. Дальнейшее развитие исследования целесообразно связать с расширенной апробацией модели на реальных данных разных образовательных программ, расширением её до многомерного случая с несколькими компетенциями, учётом эффектов дисциплин, преподавателей и оценочных средств, а также с интеграцией модели в цифровой профиль или цифровой двойник образовательной траектории обучающегося.

**Конфликт интересов**

Не указан.

Рецензия

Все статьи проходят рецензирование. Но рецензент или автор статьи предпочли не публиковать рецензию к этой статье в открытом доступе. Рецензия может быть предоставлена компетентным органам по запросу.

Conflict of Interest

None declared.

Review

All articles are peer-reviewed. But the reviewer or the author of the article chose not to publish a review of this article in the public domain. The review can be provided to the competent authorities upon request.

Список литературы / References

1. Баженова И.В. Интеллектуальная модель оценки уровня расчётно-алгоритмического компонента вычислительного мышления обучающихся / И.В. Баженова, М.М. Клунникова, Н.И. Пак // Информатика и образование. — 2022. — Т. 37, № 4. — С. 71–79. — DOI: 10.32517/0234-0453-2022-37-4-71-79.
2. Samejima F. Estimation of latent ability using a response pattern of graded scores / F. Samejima // Psychometrika Monograph Supplement. — 1969. — № 17. — DOI: 10.1007/BF03372160.
3. Wang C. On longitudinal item response theory models: a didactic / C. Wang, S.W. Nydick // Journal of Educational and Behavioral Statistics. — 2020. — Vol. 45, № 3. — P. 339–368. — DOI: 10.3102/1076998619882026.
4. Wang X. Bayesian analysis of dynamic item response models in educational testing / X. Wang, J.O. Berger, D.S. Burdick // The Annals of Applied Statistics. — 2013. — Vol. 7, № 1. — P. 126–153. — DOI: 10.1214/12-AOAS608.
5. Tong C. Deep knowledge tracing and cognitive load estimation for personalized learning path generation using neural network architecture / C. Tong, C. Ren // Scientific Reports. — 2025. — Vol. 15. — DOI: 10.1038/s41598-025-10497-x.
6. Bai Y. A survey of explainable knowledge tracing / Y. Bai, J. Zhao, T. Wei [et al.] // Applied Intelligence. — 2024. — Vol. 54, № 8. — P. 6483–6514. — DOI: 10.1007/s10489-024-05509-8.
7. Kabashkin I. AI-based digital twins of students: a new paradigm for competency-oriented learning transformation / I. Kabashkin // Information. — 2025. — Vol. 16. — DOI: 10.3390/info16100846.
8. Tesio L. Interpreting results from Rasch analysis. Advanced model applications and the data-model fit assessment / L. Tesio, A. Caronni, D. Kumbhare [et al.] // Disability and Rehabilitation. — 2024. — Vol. 46, № 3. — P. 604–617. — DOI: 10.1080/09638288.2023.2169772.
9. Братищенко В.В. Модель с латентными параметрами для пошаговой процедуры оценивания результатов обучения / В.В. Братищенко // Моделирование, оптимизация и информационные технологии. — 2022. — Т. 10, № 1. — DOI: 10.26102/2310-6018/2022.36.1.015.
10. Lee M. Evaluation of residual-based fit statistics for item response theory models in the presence of non-responses / M. Lee, J. Jung // Educational and Psychological Measurement. — 2025. — Vol. 86, № 3. — P. 602–626. — DOI: 10.1177/00131644251393444.
11. Fujimoto K.A. The accuracy of Bayesian model fit indices in selecting among multidimensional item response theory models / K.A. Fujimoto, C.F. Falk // Educational and Psychological Measurement. — 2024. — Vol. 84, № 2. — P. 217–244. — DOI: 10.1177/00131644231165520.

Список литературы на английском языке / References in English

1. Bazhenova I.V. Intellektual'naya model' otsenki urovnya raschetno-algoritmicheskogo komponenta vychislitel'nogo myshleniya obuchayushchikhsya [An intelligent model for assessing the level of the computational-algorithmic component of students' computational thinking] / I.V. Bazhenova, M.M. Klunnikova, N.I. Pak // Informatika i obrazovaniye [Informatics and Education]. — 2022. — Vol. 37, № 4. — P. 71–79. — DOI: 10.32517/0234-0453-2022-37-4-71-79. [in Russian]
2. Samejima F. Estimation of latent ability using a response pattern of graded scores / F. Samejima // Psychometrika Monograph Supplement. — 1969. — № 17. — DOI: 10.1007/BF03372160.
3. Wang C. On longitudinal item response theory models: a didactic / C. Wang, S.W. Nydick // Journal of Educational and Behavioral Statistics. — 2020. — Vol. 45, № 3. — P. 339–368. — DOI: 10.3102/1076998619882026.
4. Wang X. Bayesian analysis of dynamic item response models in educational testing / X. Wang, J.O. Berger, D.S. Burdick // The Annals of Applied Statistics. — 2013. — Vol. 7, № 1. — P. 126–153. — DOI: 10.1214/12-AOAS608.
5. Tong C. Deep knowledge tracing and cognitive load estimation for personalized learning path generation using neural network architecture / C. Tong, C. Ren // Scientific Reports. — 2025. — Vol. 15. — DOI: 10.1038/s41598-025-10497-x.
6. Bai Y. A survey of explainable knowledge tracing / Y. Bai, J. Zhao, T. Wei [et al.] // Applied Intelligence. — 2024. — Vol. 54, № 8. — P. 6483–6514. — DOI: 10.1007/s10489-024-05509-8.
7. Kabashkin I. AI-based digital twins of students: a new paradigm for competency-oriented learning transformation / I. Kabashkin // Information. — 2025. — Vol. 16. — DOI: 10.3390/info16100846.
8. Tesio L. Interpreting results from Rasch analysis. Advanced model applications and the data-model fit assessment / L. Tesio, A. Caronni, D. Kumbhare [et al.] // Disability and Rehabilitation. — 2024. — Vol. 46, № 3. — P. 604–617. — DOI: 10.1080/09638288.2023.2169772.
9. Bratishchenko V.V. Model' s latentnymi parametrami dlya poshagovoy protsedury otsenivaniya rezul'tatov obucheniya [A model with latent parameters for a step-by-step procedure for assessing learning outcomes] / V.V. Bratishchenko // Modelirovaniye, optimizatsiya i informatsionnyye tekhnologii [Modeling, Optimization and Information Technology]. — 2022. — Vol. 10, № 1. — DOI: 10.26102/2310-6018/2022.36.1.015. [in Russian]



10. Lee M. Evaluation of residual-based fit statistics for item response theory models in the presence of non-responses / M. Lee, J. Jung // Educational and Psychological Measurement. — 2025. — Vol. 86, № 3. — P. 602–626. — DOI: 10.1177/00131644251393444.
11. Fujimoto K.A. The accuracy of Bayesian model fit indices in selecting among multidimensional item response theory models / K.A. Fujimoto, C.F. Falk // Educational and Psychological Measurement. — 2024. — Vol. 84, № 2. — P. 217–244. — DOI: 10.1177/00131644231165520.