

# ИССЛЕДОВАНИЕ ТОЧНОСТИ ИЗМЕРЕНИЯ ЛАТЕНТНОЙ ПЕРЕМЕННОЙ В ЗАВИСИМОСТИ ОТ ЗАШУМЛЕНИЯ РЕЗУЛЬТАТОВ ТЕСТИРОВАНИЯ

## *Investigation of Measurement Precision of Latent Variables Depending on the Noise Test Results*

**Anatoly Maslak**

Kuban State University, Russian Federation

**Abstract.** *Measurement precision is an important aspect of experimental studies because it affects the effectiveness of decisions. The precision of measurements is influenced by many factors, including the quality of the original data or how the data was obtained. When measuring a latent variable, one of these factors is the precision of the indicators. For example, there may be situations in which, for various reasons, the subject gave the wrong answer instead of the correct answer and vice versa. This fully applies to conducting surveys. In general, we are talking about the noise of the original data. The aim of the work is to analyze the precision of measurement of the latent variable depending on the degree of noise of the original data. The results of the simulation experiment are used as initial data, which allows to simulate the necessary experimental situations. The analysis is conducted within the framework of the theory of latent variables, which allows to obtain estimates of the latent variable on a linear scale. The measurement precision of the latent variable is determined on the basis of absolute error. The noise level of the original data varies from 0 to 0.10 %. It is shown that even minor noise of the original data significantly reduces the precision of the measurement of the latent variable.*

**Keywords:** *latent variable, precision of measurement, Rasch model.*

### **Введение Introduction**

Измерение латентных переменных является важным атрибутом экспериментального исследования. Во-первых, уточняется сам смысл измеряемой латентной переменной, поскольку латентная переменная определяется в виде набора индикаторов, то есть в виде набора ее проявлений. Во-вторых, латентная переменная измеряется на линейной шкале, что позволяет использовать широкий класс статистических процедур, влияющих на латентную переменную (Rasch, 1980). Вот почему теория латентных переменных находит широкое применение в самых различных социальных

системах (Bond, 2015; Engelhard, 2013; Krabe, 2017; Leus & Maslak, 2018; Maslak et al. 2005; Maslak et al. 2018; Maslak & Pozdniakov, 2018; Maslak et al. 2015).

Поскольку процедура формирования набора индикаторов является неформальной и зависит от предпочтений того или иного специалиста, то возникают вопросы, а какие должны быть формальные параметры набора индикаторов, чтобы обеспечить наибольшую точность измерения латентной переменной. В лаборатории объективных измерений Кубанского государственного университета проведены многие исследования по оценке влияния параметров набора индикаторов на точность измерения латентной переменной. Необходимо отметить работы, в которых исследуется точность измерения латентной переменной в зависимости от диапазона варьирования индикаторов (Маслак et al., 2017) и в зависимости от части правильных ответов в задачах тестирования (Маслак et al., 2012).

Целью исследования является точность измерения латентной переменной в зависимости от числа ошибок в дихотомических индикаторах. На практике это означает, например, что на некоторые тестовые задания по некоторым причинам испытуемый дал неправильные ответы. Очевидно, что чем больше ошибок, тем ниже точность вычисления латентной переменной. Это обуславливает необходимость определения точности измерения латентной переменной в зависимости от числа ошибок в значениях индикаторов. Фактически речь идет о зашумлении данных. Использовались одиннадцать уровней зашумления полученных наборов: 0,00 %, 0,01 %, 0,02 %, ..., 0,10 %. Процедура зашумления состояла в том, что значение случайным образом выбранного элемента набора заменялось на противоположное, то есть «1» заменялась на «0», а «0» на «1». Ошибка измерения вычислялась как абсолютная разность между моделируемым значением и соответствующей оценкой.

Поскольку шкала измерения является интервальной, то исследуется абсолютная ошибка измерений. Поэтому необходимо установить количественную зависимость абсолютной ошибки измерения от числа ошибок в значениях индикаторов и подготовить соответствующие рекомендации.

Исследование проводится на основе имитационного эксперимента в двукратной повторности.

## **Данные** *Data*

Данные имитационного эксперимента генерируются в соответствии с моделью измерения (моделью Раша), т.е. априорно сгенерированные данные адекватны модели измерения.

Для генерации матрицы тестирования использовалась следующая схема. Было выбрано 30 студентов, уровень подготовленности которых равномерно варьировался на отрезке  $[-4; +4]$  в логитах. Этот диапазон практически охватывает большинство практических задач. Число индикаторов было выбрано равным 30, которые также равномерно варьировались в диапазоне  $[-5; +5]$  логит. Эта ситуация характерна для многих практических ситуаций.

Модель Раша для измерения уровня знаний испытуемых имеет вид

$$P_{ij} = \frac{e^{\beta_i - \delta_j}}{1 + e^{\beta_i - \delta_j}}, \quad (1)$$

где  $\beta_i$  – уровень подготовленности  $i$ -ого испытуемого (в логитах),  
 $\delta_j$  – трудность  $j$ -ого задания (в логитах).

Здесь уровень подготовленности и трудность тестового задания являются латентными переменными. На основе матрицы результатов тестирования размера  $I \times J$  (всего  $I \times J$  данных) производится измерение латентных переменных (всего  $I + J$  параметров). Имитационный эксперимент проводился в двукратном повторении.

### Методология исследования Research Methodology

Процедура имитационного моделирования состоит в следующем. Прежде всего, на основе модели Раша для дихотомических индикаторов по формуле (1) вычисляется  $P_{ij}$  – вероятность правильного ответа  $i$ -ого испытуемого на  $j$ -ое задание. Затем на основе вычисленных вероятностей по формуле (1) генерируются результаты матрицы тестирования:

$$X_{ij} = \text{Int} (P_{ij} - \text{Rnd} + 1), \quad (2)$$

где  $\text{Int} (Y)$  – целая часть числа  $Y$ ,  
 $\text{Rnd}$  – случайное число, равномерно распределенное в интервале  $(0; 1)$ .

В табл. 1 в качестве примера приведен фрагмент матрица результатов тестирования.

Таблица 1. Матрица результатов тестирования, число инвертированных значений индикаторов равно 0 %

Table 1 Matrix of test results, the percentage of inverted indicator values is 0 %

№ п/п	Индикаторы
-4.000	11010000000000000000000000000000
-3.724	11101000000000000000000000000000
-3.448	01111101010000000000000000000000
-3.172	10100110100000000001000000000000
-2.897	11111101010010000000000000000000
-2.621	11101011001000000000000000000000
-2.345	11111110100100000000000000000000
-2.069	11111011110000000001000000000000
-1.793	11101111011000000100000000000000
-1.517	11111101011110100000000000000000
-1.241	11111111110000000000000000000000
-0.966	11111111111110100000000000000000
-0.690	11111110011100001000000000000000
-0.414	11111101011110111000000000000000
-0.138	111110101011100110000001000000
0.138	111111111111100000100100000000
0.414	111111111111110000001000000000
0.690	111111111111110100100000000000
0.966	111111111101001111000000000000
1.241	1111111111111110101100000000
1.517	111111111111110110011000000000
1.793	11111111111110111001110001000
2.069	111111111111111011010000001
2.345	111111111111111101001000001
2.621	1111111111111011111010010000
2.897	1111111111111111111011010010
3.172	1111111111111111111100001000
3.448	111111111111111111111100000
3.724	1111111111111101100101110000
4.000	11111111111111111111110001100

Уровень значимости статистики Хи-квадрат (Chi-square probability) для этой матрицы равна 0,767, индекс дифференциации объектов (person separation index) равен 0,939.

**Таблица 2. Матрица результатов тестирования, число инвертированных значений индикаторов равно 0,10 %**

**Table 2 Matrix of test results, the percentage of inverted indicator values is 0.10 %**

№ п/п	Индикаторы
-4.000	101010111001000100000000000000
-3.724	101010000100100000000000000000
-3.448	010111000100000000000000000100
-3.172	111110010100000001000000000000
-2.897	100101111000000000010000000000
-2.621	111101011000100000000001001000
-2.345	111011101000000000000010000000
-2.069	100101100011000010000000000000
-1.793	111100111101000000000001000000
-1.517	111111000101000000010000000000
-1.241	111111110110000010000100001010
-0.966	111111100110100101010000000000
-0.690	111111011110010000100000000000
-0.414	111111010101010010010100000000
-0.138	011101111111111010001111000100
0.138	111111110111110001000100000000
0.414	111111111111111000000100010110
0.690	111111111111101111000010100000
0.966	111111011110101110101110010000
1.241	110011111101011111000000000000
1.517	111011111111111110011000010000
1.793	111101111111111011010000000000
2.069	11011111111011100110000010011
2.345	011111111111111111111111001000
2.621	111111101110111011110101000000
2.897	111111111111111111101100001000
3.172	1111011011101110011111111010010
3.448	101111111110101111111100101100
3.724	11111111111111111111101100010
4.000	0111111101111111011111111100000

Уровень значимости статистики Хи-квадрат (Chi-square probability) для этой матрицы равна 0,258, индекс дифференциации объектов (person separation index) равен 0,865.

## Результаты Results

Далее сгенерированные данные использовались для вычисления оценок  $\hat{\beta}_i$  и  $\hat{\delta}_j$  истинных значений латентных переменных  $\beta_i$  и  $\delta_j$ . Для нахождения оценок использовалась лицензионная диалоговая система ИЛП (Измерение Латентных Переменных), разработанная в лаборатории объективных измерений Кубанского государственного университета.

Затем вычислялась абсолютная погрешность вычисления латентной переменной:

$$\Delta i = |\hat{\beta}_i - \beta_i| \quad (3)$$

Число ошибок в индикаторах варьируется на 11 уровнях – 0,00 %, 0,01%, 0,02 %, ..., 0,10 %. Имитационный эксперимент был проведен дважды, соответствующие результаты представлены в табл. 2.

Таблица 3. Точность измерения латентной переменной в зависимости от зашумления данных  
Table 3 Precision of measurement of latent variable depending on data noise

	Доля ошибок										
	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09	0,10
Повторение 1	0,45	0,45	0,52	0,66	0,81	0,74	0,85	0,86	1,18	1,15	1,06
Повторение 2	0,47	0,56	0,61	0,86	0,79	1,01	0,85	0,92	0,99	1,12	1,14
Среднее	0,46	0,51	0,57	0,76	0,80	0,88	0,85	0,89	1,09	1,14	1,10

Иллюстрация полученных результатов представлена на рис. 1.

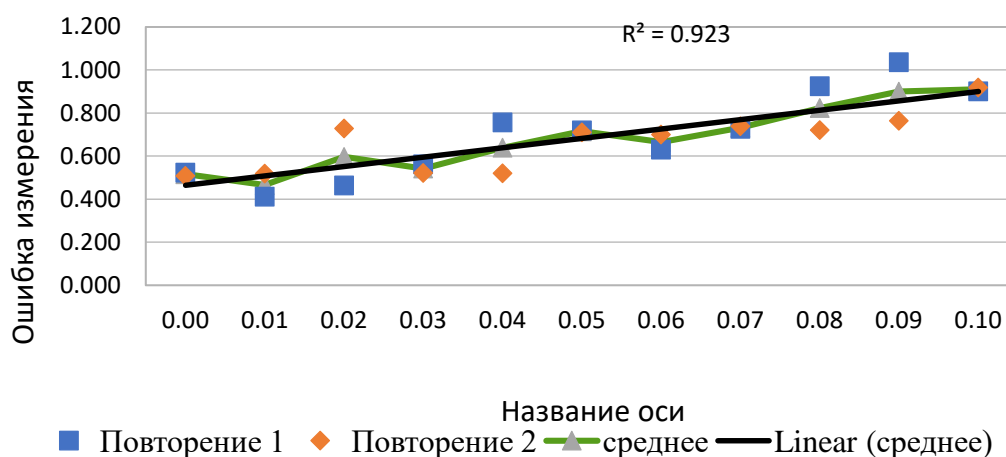


Рисунок 1. Абсолютная ошибка измерения латентной переменной в зависимости от зашумления результатов тестирования  
Figure 1 Absolute error of measurement of the latent variable depending on the noise of the test results

Как следует из рис. 1 абсолютная ошибка измерения латентной переменной прямо пропорционально зависит от доли ошибок в индикаторах. Необходимо отметить, что коэффициент детерминации очень высокий и равен 0,923. Минимальная ошибка наблюдается в случае отсутствия ошибок в индикаторах и равна 0,5 логит, максимальная ошибка наблюдается, когда число ошибочных значений равно 0,1 % и равна 0,9 логит.

В целом, наличие ошибок в индикаторах существенно влияет на точность измерения, поэтому таких ошибок необходимо избегать.

### **Выводы** *Conclusions*

Проведенное исследование показало, что зашумление данных тестирования и/или опросов в значительной степени увеличивает ошибку измерения. Ошибка измерения латентной переменной прямо пропорционально зависит от степени зашумления результатов тестирования и/или опросов. Коэффициент линейной детерминации очень высок и равен 0,923. При степени зашумления данных в 0,10 % абсолютная ошибка измерения увеличивается на 0,5 логит. Это свидетельствует о необходимости тщательной организации проведения процедуры тестирования и/или опроса для устранения причин зашумления данных.

### **Summary**

A simulation experiment was carried out in two-fold repetition to generate sets of test results with the specified parameters. Eleven noise levels of the obtained sets were used: 0.00 %, 0.01 %, 0.02 %, ..., 0.10 %. The noise procedure was the following: the value of a randomly selected element of the set was inverted; that is, "1" was replaced by "0" and "0" by "1". The investigation was conducted within the framework of the theory of latent variables. To estimate the latent variable, the interactive system MLV (Measurement of Latent Variables) developed in the laboratory of objective measurements of the Kuban State University (Russia) was used. The measurement error was calculated as the absolute difference between the simulated value and the corresponding estimate. It is shown that the measurement error is directly proportional to the degree of data noise; the coefficient of determination is 0.923. This means that testing and polling procedures must be carefully organized to eliminate data noise.

## Литература References

- Bond, Trevor, G. (2015). *Applying the Rasch model: fundamental measurement in the human sciences*. Third edition. Trevor G. Bond & Christine M. Fox. Routledge. New York. NY.
- Engelhard, G. (2013). *Invariant Measurement: Using Rasch Models in the Social, Behavioral, and Health Sciences*. Routledge, New York, NY.
- Krabe, P. (2017). *The Measurement of Health and Health Status: Concepts, Methods, and Applications from a Multidisciplinary Perspective*. Academic Press, Elsevier. London.
- Leus, O., & Maslak, A. (2018). Measurement and Analysis of Teachers' Professional Performance. *Society. Integration. Education. Proceeding of the International Scientific Conference, II*, 308–319.
- Maslak, A. A., Karabatsos, G., Anisimova, T. S., & Osipov, S. A. (2005). Measuring and Comparing Higher Education Quality between Countries Worldwide. *Journal of Applied Measurement*, 6(4), 432 – 442.
- Maslak, A., Moiseev, S., & Nasonova, T. (2018). Application of the Rasch Method of Evaluating Latent Variables in Management and Administration. *Society. Integration. Education. Proceeding of the International Scientific Conference, VI*, 330–340.
- Maslak, A., & Pozdnyakov, S. (2018). Measurement and Multifactorial Analysis of Students' Patriotism. *Society. Integration. Education. Proceeding of the International Scientific Conference, I*, 373–383.
- Maslak, A. A., Rybkin, A. D., Anisimova, T. S., & Pozdnyakov, S. A. (2015). Monitoring of pupils' imagination within the framework of creativity formation program. *Mediterranean Journal of Social Sciences*, 6(6), Supplement 5, Special edition, 234-241.
- Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests (Expanded edition, with foreword and afterword by Benjamin D. Wright)*. Chicago: University of Chicago Press.
- Маслак, А. А., Моисеев, С. И., Осипов, С. А., & Поздняков, С. А. (2017). Исследование точности измерения латентной переменной в зависимости от диапазона варьирования набора индикаторов. *Радіоелектроніка, інформатика, управління. 1 (40)*, 42-49.
- Маслак, А. А., Осипов, С. А., Алиакпарова, М. М., & Филист, С. А. (2012). Оценка влияния частично правильных ответов учащихся на точность измерения уровня их подготовленности. *Известия Юго-Западного государственного университета. Серия: Управление, вычислительная техника, информатика. Медицинское приборостроение*, 2-2, 76-87.