

Алгоритмы обработки результатов

На основе первичной информации, полученной после сканирования, распознавания и верификации бумажных бланков или программной обработки данных, полученных из центров сканирования, формируется первичная матрица (таблица) результатов тестирования, строки которой соответствуют участникам оценки, а столбцы – заданиям оценки.

С помощью полученной матрицы вычисляется первичный балл каждого участника оценки и индекс сложности каждого задания (% участников, верно выполнивших данное задание). Эти показатели ложатся в основу дальнейшего статистического анализа результатов.

Дальнейший статистический анализ¹ состоит из трех этапов.

На первом этапе производится классическая (первичная) обработка результатов оценки, в ходе которой вычисляются тестологические характеристики каждого задания в данном предъявлении варианта теста, а также показатели надежности и валидности теста. На этом же этапе осуществляется обработка всеоров ответов к заданиям с кратким ответом. Первичная обработка предшествует вторичной и предназначена для разработчиков заданий и тестологов. На основе данной информации можно сделать выводы о выборке, качестве заданий вариантов теста и в целом о вариантах оценки.

Вторичная обработка предполагает вычисление показателей, связанных с оценкой уровня подготовки каждого участника, распределением первичных баллов, процентами выполнения заданий и т.п. Файлы, полученные в ходе вторичной обработки, предназначены для участников экзамена и административно-территориальных единиц, в которых проводился экзамен. Для выдачи отчетов вторичной обработки в дереве административно-территориальных единиц выделяется до 4 основных уровней административно-

-
1. Балыхина Т.М. Словарь терминов и понятий тестологии. – М: МГУП, 2000. – 161 с.
 2. Нейман Ю.М. Хлебников В.А. Введение в теорию моделирования и параметризации педагогических тестов. Москва: Прометей, 2000.
 3. Рэш Дж. Индивидуальный подход к анализу вопросов. – В кн. «Математические методы в социальных науках» - М.: Прогресс, 1973, стр. 91- 116.
 4. Хлебников В.А. (редактор) Развитие системы тестирования в России. – Материалы ежегодной Всероссийской конференции. – М: ЦТМО, 2001-2003.
 5. Чельшкова М.Б. Разработка педагогических тестов на основе современных математических моделей. – М: МИСИС, 1995.
 6. Шмелев А.Г. (редактор) Тесты для старшеклассников и абитуриентов. Телетестинг.– Москва: Первое сентября, 2000. – 132с.
 7. Cambridge International Examination. - Cambridge (UK): Local Examinations Syndicate, 2000. – 61 p.
 8. Gronlund N.E., Linn R.L. Measurement and Evaluation in Teaching. 6th edition.– N.Y.-L.: Macmillan, 1990. – 525 p.
 9. Manual for Scholastic Aptitude Test. – Princeton (N.J.): ETS, 1998.
 10. Standards for educational and psychological testing. – Washington: American Psychological Association, 1974.

территориальных единиц – вся выборка, регионы или группы регионов и участники экзамена.

На третьем этапе в результате обработки данных формируются показатели, используемые для анализа качества заданий и вариантов оценки и интерпретации результатов оценки.

В результате обработки данных формируются следующие показатели, получаемые для анализа и интерпретации результатов оценочной процедуры.

Показатели первичной обработки:

Показатели по заданиям:

1. Номер задания в варианте
2. Название задания в варианте
3. Количество участников, решавших данное задание
4. Количество участников, не решавших данное задание
5. Процент участников, не решавших данное задание
6. Количество участников, не приступивших к данному заданию
7. Процент участников, не приступивших к данному заданию
8. Точечно-бисериальный коэффициент (ТБКК) задания
9. Процент выполнения слабой группой (27% с наименьшим рейтингом)
10. Процент выполнения сильной группой (27% с наибольшим рейтингом)
11. Коэффициент дискриминативности
12. Распределение участников по дистракторам задания с выбором ответа или баллам политомического задания
13. ТБКК по дистракторам задания с выбором ответа или баллам политомического задания
14. ТБКК для участников, не приступивших к заданию
15. Уровень сложности
16. Проблемы задания по кодификатору проблем
17. В кодификаторе проблем вынесены проблемы по проценту выполнения, несоответствующему уровню заданий, и проблемы заданий, выявленные по корреляционным коэффициентам (ТБКК).
18. Группа проблем по кодификатору групп проблем
19. Процент выполнения задания
20. Процент выполнения задания процентильными группами
21. Дата проведения процедуры

Показатели по вариантам:

1. Номер варианта
2. Количество участников, выполнявших данный вариант
3. Средний процент выполнения варианта
4. Максимальный балл варианта
5. Средний балл
6. Стандартное отклонение первичного балла
7. Дисперсия первичного балла
8. Ассиметрия первичного балла
9. Эксцесс первичного балла
10. Средняя пятибалльная отметка
11. Стандартное отклонение отметки
12. Коэффициент надежности варианта
13. Распределение участников по первичным баллам

Данные интегральные характеристики дают показатели по серии вариантов и по всей выборке участников.

По всей выборке для отчетов вычисляются следующие показатели:

1. Количество участников
2. Распределение первичных баллов

Показатели первичной обработки сохраняются в файл csv для возможности импорта показателей в базу тестовых заданий и анализа качества заданий, вариантов, выборки.

Показатели вторичной обработки:

Показатели вторичной обработки, участвующие в отчетах *для административно – территориальных единиц*:

1. Количество участников
2. Распределение участников по первичному баллу
3. Процент выполнения каждого тестового задания
4. Процент выполнения каждого тестового задания группами участников

Показатели вторичной обработки, участвующие в отчетах *для участников экзамена*:

1. Вариант теста
2. Баллы за задание
3. Первичный балл

Виды отчетов

При создании статистических отчетов по исследованию можно выделить 4 группы форм:

– для оценки состояния системы образования административно-территориальной единицы,

– для оценки индивидуальных результатов участников оценки,

– для использования при разработке вариантов оценки.

Виды отчетов первичной обработки:

1. Статистические данные регионов
2. Общие статистические данные заданий по вариантам оценки
3. Структура заданий по иерархии проблем (карманам)
4. Сравнительный анализ процента выполнения заданий по группам баллов
5. Сравнительный анализ процента выполнения заданий по группам баллов по

вариантам оценки

6. Сравнительный анализ по процентилям по вариантам оценки

7. Распределение первичных баллов

Виды отчетов вторичной обработки

Для региона предусмотрены следующие отчеты по участникам:

1. Таблица результатов участников (по рейтингу или по алфавиту)
2. Индивидуальные результаты участников

Для различных территориальных единиц предусмотрены следующие отчеты:

1. Индивидуальные результаты участников, могут использоваться для дальнейшего импорта в БД

2. Выполнение заданий
3. Распределение первичных баллов
4. Выполнение заданий группами участников

Для всей выборки отчет содержит показатели всей выборки, а также вложенных территориальных единиц.

Для региона отчет содержит показатели региона, а также вложенных территориальных единиц.

Анализ распределений по субъектам Российской Федерации показывает, насколько однородны региональные образовательные системы, а анализ распределений по ОО внутри субъекта Российской Федерации показывает, насколько однородны образовательные системы внутри региона.

В формах статистических отчетов выделены оцениваемые параметры, описаны типовые критерии оценивания, информация о результатах оценивания визуализирована при представлении форм статистики.

Показатели для анализа качества вариантов КИМ

Для анализа качества вариантов анализируются следующие показатели: средний процент выполнения и средний первичный балл, стандартное отклонение первичного балла, асимметрия, эксцесс и надежность.

– *Средний процент выполнения и средний первичный балл* показывают насколько были сбалансированы по трудности варианты оценки. Но необходимо учитывать специфику выборки, не всегда большое расхождение по средним баллам говорит о том, что варианты неравнозначны по трудности, возможно, это особенности регионов, в которых они выполнялись.

– *Стандартное отклонение* равно корню квадратному из дисперсии. Чем шире распределение, тем больше значение этого показателя, а значит, тем более информативен вариант теста.

– *Асимметрия* распределения показывает сбалансированность теста, если ее значение близко к 0, то тест сбалансирован по трудности заданий, если ее значение отрицательно, то это говорит о преобладании легких заданий. Чем больше значение асимметрии, тем более сложные задания включены в тест, то есть основная часть значений индивидуальных баллов лежит слева от среднего значения, что характерно для трудных заданий, требуется корректировка содержания по трудности.

– *Значения эксцесса* позволяют получить представление о том, являются ли полигон частот распределения или гистограмма, островершинной или плоской. Для нормального распределения эксцесс должен быть нулевым. Если значения эксцесса больше 2, то это говорит об островершинной кривой, это еще раз позволяет подтвердить, что есть проблемы с трудностью заданий.