

Линда Крокер, Джеймс Алгина

Введение в классическую и современную теорию тестов

Перевод на русский язык
под общей редакцией В.И. Звонникова и М.Б. Челышковой



Москва
Логос
2010

Перевод и издание на русском языке книги Л. Крокер и Дж. Алгины
«Введение в классическую и современную теорию тестов»
осуществлены при поддержке Американских советов
по международному образованию

УДК 159.9.07
ББК 74.26в6
К83

Крокер Л.

К83 Введение в классическую и современную теорию тестов: учебник / Л. Крокер, Дж. Алгина; под общей ред. В.И. Звонникова и М.Б. Чельшковой. — М.: Логос, 2010. — 668 с.

ISBN 978-5-98704-437-5

Раскрыта сущность теории тестов, показаны преемственность и отличия ее классической и современной версий. Изложено введение в теорию измерений. Представлены статистические концепции в теории тестирования. Дано введение в шкалирование, показан процесс конструирования тестов. Раскрыто понятие надежности в контексте классической концепции истинной оценки и разработки процедур оценивания справедливости тестов. Освещена проблема валидности тестов и изложены пути ее решения. Охарактеризованы тестовые оценки и методы их интерпретации. Каждая глава содержит краткие выводы и упражнения на закрепление полученного материала и проверку знаний.

Для студентов высших учебных заведений, получающих образование по педагогическим специальностям и направлениям, а также для слушателей учреждений повышения квалификации и переподготовки учителей и преподавателей. Может использоваться в учебном процессе вузов при подготовке социологов, психологов, экономистов, представителей других областей социально-гуманитарных наук, знания, применяющих тесты для измерения и оценки человеческих способностей, знаний и качеств, а также явлений и процессов общественной и культурной жизни. Представляет особый интерес для ученых и специалистов, занимающихся проблемами измерений.

УДК 159.9.07
ББК 74.26в6

ISBN 978-5-98704-437-5

© Cengage Learning, 2006, 2008
© Найденова Н.Н., Симкин В.Н.,
Чельшкова М.Б., перевод, 2010
© Звонников В.И., Чельшкова М.Б.,
предисловие к русск. изд., 2010
© Логос, 2010

ОГЛАВЛЕНИЕ

Предисловие к русскому изданию	11
Предисловие	17
Раздел I. Введение в теорию измерений	21
Глава 1. Что такое теория тестов?	23
Проблемы измерений психологических конструктов.....	26
Теория тестов как дисциплина.....	28
Структура данного учебника	34
<i>Выводы</i>	36
<i>Упражнения</i>	36
Глава 2. Статистические концепции в теории тестов	38
Тестовые оценки как дискретные переменные для конечных популяций.....	38
Описание индивидуальных результатов с помощью z-оценок.....	46
Бесконечные популяции и непрерывные переменные.....	47
Нормальное распределение	48
Описание связи между двумя переменными	54
Предсказание индивидуальных возможностей выполнения теста.....	64
<i>Выводы</i>	71
<i>Упражнения</i>	73
Глава 3. Введение в шкалирование	77
Вещественные числа и шкалы измерения	77
Уровни измерительных шкал	79
Подходы к шкалированию в разработке тестов	83
Уровни шкал для субъектно-центрированных измерений.....	97
<i>Выводы</i>	101
<i>Упражнения</i>	102

Глава 4. Процесс конструирования теста	105
Определение целей применения тестовой оценки	106
Определение характеристик поведения для отображения конструкта	107
Построение выборки содержания	108
Подготовка спецификаций тестов	113
Конструирование заданий	117
Рецензирование	125
Предварительная апробация заданий	127
Следующие шаги	128
<i>Выводы</i>	128
<i>Упражнения</i>	130
Глава 5. Тестовые оценки как составные	133
Схемы оценивания заданий	134
Дескриптивная (описательная) статистика для недихотомических переменных	135
Дескриптивная статистика для дихотомических переменных	137
Дисперсия составной оценки	141
Практические приложения для конструирования теста	145
<i>Выводы</i>	148
<i>Упражнения</i>	149
Раздел II. Надежность	151
Глава 6. Надежность и классическая модель истинной оценки	153
Классическая модель истинной оценки	155
Показатель надежности и коэффициент надежности	164
Надежность составных тестовых оценок	168
Стандартная ошибка измерения	174
Альтернативные определения истинной оценки и ошибки	177
<i>Выводы</i>	181
<i>Упражнения</i>	183
Глава 7. Процедуры для оценивания надежности	186
Процедуры, требующие двух предъявлений теста	187
Методы, требующие одноразового предъявления теста	191
Межрейтерная надежность	203
Факторы, которые влияют на коэффициенты надежности	204

Приближенное вычисление истинных оценок	208
Надежность разности оценок.....	210
Использование оценок ошибки в интерпретации тестового балла.....	213
Сообщение данных по надежности	216
<i>Выводы</i>	217
<i>Упражнения</i>	218
Глава 8. Введение в теорию генерализации	223
Г-исследование и Р-исследование	224
Введение коэффициентов генерализации для однофасетных дизайнов.....	227
Вложенный однофасетный дизайн Г-исследования (дизайн 4)	241
Генеральные совокупности с фиксированными фасетами.....	243
Использование теории генерализации с данными, отличающимися от субъективных оценок экспертов	245
Стандартные ошибки измерения для абсолютных и относительных решений	246
Теория генерализации для двухфасетного дизайна	251
<i>Выводы</i>	265
<i>Упражнения</i>	266
Глава 9. Коэффициенты надежности для критериально- ориентированных тестов	269
Использование критериально-ориентированных измерений.....	270
Теория надежности для приближенного оценивания домен-оценок.....	271
Теория надежности для классификаций по категориям мастерства.....	275
Точность решения	292
<i>Выводы</i>	294
<i>Упражнения</i>	296
Раздел III. Валидность.....	299
Глава 10. Введение в валидность.....	301
Содержательная валидность	302
Критериально-ориентированная валидизация	310
Конструктная валидность	319
Пересечение подходов к валидизации	326
Коэффициенты валидности для истинных оценок.....	327
<i>Выводы</i>	329
<i>Упражнения</i>	331

Глава 11. Статистические процедуры для прогнозирования и классификации	335
Частная корреляция	336
Множественная регрессия	339
Дискриминантный анализ	352
<i>Выводы</i>	360
<i>Упражнения</i>	361
Глава 12. Систематическая ошибка отбора	365
Основные концепции и понятия	366
Группы большинства и меньшинства	370
Психометрические определения систематической ошибки	370
Справедливые процедуры отбора	374
Критика моделей	379
Подход к отбору, связанный с теорией решений	380
<i>Выводы</i>	387
<i>Упражнения</i>	388
Глава 13. Факторный анализ	390
Пример с гипотетическими данными	391
Факторы и факторные нагрузки	391
Вращение	393
Коррелированные факторы	396
Число факторов	398
Модель факторного анализа	399
Общность и уникальность	400
Пример с реальными данными	401
Исследовательский и подтверждающий факторный анализ	410
<i>Выводы</i>	412
<i>Упражнения</i>	412
Раздел IV. Анализ заданий в разработке теста	415
Глава 14. Анализ заданий	417
Трудность задания, среднее значение и дисперсия	418
Дискриминативность (различительная способность) задания	420
Надежность заданий и показатели валидности	428
Проведение исследования при анализе заданий	430

Анализ заданий для критериально-ориентированных тестов.....	441
<i>Выводы</i>	449
<i>Упражнения</i>	451
Глава 15. Введение в современную теорию тестов	454
Базовые понятия теории IRT	456
Кривая интегральной функции нормального распределения (нормальная огива)	463
Связь теории IRT с классической теорией тестов.....	468
Логистические модели	471
Приближенное оценивание параметров	473
Выбор модели	482
Приложения теории IRT	484
<i>Выводы</i>	496
<i>Упражнения</i>	496
Глава 16. Обнаружение заданий, приводящих к систематическим ошибкам	502
Методы, основанные на современной теории тестов (IRT)	504
Методы хи-квадрат	512
Преимущества и неудобства различных методов	524
<i>Выводы</i>	525
<i>Упражнения</i>	526
Раздел V. Тестовые оценки и их интерпретация	529
Глава 17. Поправка на угадывание и другие методы оценивания	531
Оценивающая формула.....	532
Поощрение дополнительными вкладами в оценку (кредитами) за частичные знания.....	539
<i>Выводы</i>	540
<i>Упражнения</i>	542
Глава 18. Установление стандартов	544
Подходы к установлению стандартов	546
Эмпирическое исследование методов установления стандартов.....	553
Практические рекомендации по установлению стандартов.....	555
Технические соображения в процедуре установления стандартов.....	558
<i>Выводы</i>	568
<i>Упражнения</i>	570

Глава 19. Нормы и стандартные оценки	571
Проведение исследования нормирования	572
Вероятностная выборка	573
Описание исследования нормирования в руководстве к тесту	580
Виды нормативных оценок	582
<i>Выводы</i>	598
<i>Упражнения</i>	600
Глава 20. Выравнивание оценок по разным тестам	604
Определение выравнивания	605
Сбор данных для выравнивания	606
Выравнивание по современной теории тестов (IRT)	616
Выравнивание истинной оценки	622
Вертикальное выравнивание	624
Равенство в выравнивании	629
<i>Выводы</i>	631
<i>Упражнения</i>	632
Библиография	635
Приложение А	652
Приложение В	653

ПРЕДИСЛОВИЕ

В 1914 г. Эдвард Ли Торндайк опубликовал книгу «Теория умственных и социальных измерений», которая была признана первым учебником по теории тестов. Работа была встречена с огромным энтузиазмом коллегами-профессионалами, и в течение короткого времени изучение теории тестов стало стандартной составной частью учебного плана подготовки аспирантов в области психологии и педагогики. В течение последующих шестидесяти лет ученые расширили, уточнили и дополнили теорию, первоначально представленную Торндайком. Этот объем знаний, известный как классическая теория тестов, стал теоретической основой разработки большинства тестов для определения способностей, учебных достижений, личностных характеристик и интересов.

Однако за последние двадцать лет теория тестов стремительно продвинулась вперед, что не отмечено в традиционных учебниках. Этому продвижению во многом способствовало развитие компьютерных технологий, позволившее практически применять более сложные математические и статистические модели к обработке данных тестирования. Увеличение числа профессионалов, работающих в области психометрики, а также рост потребности объективного измерения качеств учащихся и клиентов со стороны государственных служб и образовательных программ способствовали развитию данной области. Как результат — анализ содержания последних публикаций в ведущих изданиях по педагогическим и психологическим измерениям позволяет выявить многие доминирующие темы, которые появились относительно недавно. Примеры таких тем включают методы разработки критериально-ориентированных тестов, современную теорию тестов (IRT), теорию генерализации и психометрические методы исследования систематической ошибки тестовых заданий. Прочное знание классической теории тестов становится уже недостаточным для подготовки студента или практикующего специалиста к чтению современной профессиональной литературы или к применению содержащейся в ней информации.