

Сила вероятности в психометрике

Рецензия на книгу *“Bayesian Psychometric Modelling”*¹

Ирина Угланова

Статья поступила в редакцию в мае 2023 г. Угланова Ирина Львовна — научный сотрудник Лаборатории измерения новых конструкторов и дизайна тестов Центра психометрики и измерений в образовании Института образования, Национальный исследовательский университет «Высшая школа экономики». Адрес: 101000, Москва, Потаповский пер., 16, стр. 10. E-mail: iluglanova@gmail.com. ORCID: <https://orcid.org/0000-0001-9117-5997>

Аннотация Возникновение и развитие байесовской психометрики — закономерный результат стремления психометрики уловить и уменьшить ошибку измерения. В рецензируемой книге впервые представлено систематическое описание байесовского подхода в психометрических исследованиях. Книга состоит из двух частей: в первой изложены основные положения байесовского подхода (*Foundations*), во второй — их применение в психометрическом моделировании (*Psychometrics*). Автор этой рецензии считает, что издание будет полезно тем, кто работает в частотном подходе и хотел бы узнать про байесовский. При этом тем, кто не уверен в надежности своего математического бэкграунда, автор рецензии рекомендует обращаться к дополнительным источникам — не оснащенным таким детальным математическим сопровождением. Книга не переведена на русский язык.

Ключевые слова психометрика, байесовский подход, байесовская статистика, рецензия

Для цитирования Угланова И.Л. (2023) Сила вероятности в психометрике. Рецензия на книгу *“Bayesian Psychometric Modelling”*. *Вопросы образования / Educational Studies Moscow*, № 3, сс. 231–236. <https://doi.org/10.17323/vo-2023-17952>

Power of Probability in Psychometrics

Review of the book “Bayesian Psychometric Modeling”

Irina Uglanova

Irina L. Uglanova — Researcher at the Laboratory of Measurement of New Constructs and Test Design, Centre for Psychometrics and Measurement in Education, Institute of Education, National Research University Higher School of Economics. Address: Bld. 10, 16 Potapovsky Ln, 101000 Moscow, Russian Federation. E-mail: iluglanova@gmail.com. ORCID: <https://orcid.org/0000-0001-9117-5997>

¹ Levy R., Mislevy R.J. (2016) *Bayesian Psychometric Modeling*. Boca Raton, FL: Chapman and Hall/CRC. <https://doi.org/10.1201/9781315374604>

Abstract The emergence and development of Bayesian psychometrics is a result of psychometrics' desire to reduce measurement error. This book is the first to present a systematic description of the Bayesian approach in psychometric research. The book consists of two parts: the first one presents the main principles of the Bayesian approach (Foundations), the second one includes their application in psychometric modeling (Psychometrics). The reviewer believes that the publication will be useful for those who used to work in the frequentist approach and would like to learn about the Bayesian approach. At the same time, she recommends those who are not sure of the quality of their mathematical background to additionally turn to other sources that are not equipped with such a detailed mathematical description. The book has not been translated into Russian.

Keywords psychometrics, Bayesian approach, Bayesian statistics, review

For citing Uglanova I.L. (2023) Sila veroyatnosti v psikhometrike. Retsenziya na knigu "Bayesian Psychometric Modelling" [Power of Probability in Psychometrics. Review of the book "Bayesian Psychometric Modeling"]. *Voprosy obrazovaniya / Educational Studies Moscow*, no 3, pp. 231–236. <https://doi.org/10.17323/vo-2023-17952>

*«Значительная часть этой книги посвящена тому,
как применение вероятностного подхода
к психометрическому моделированию позволяет
не только уловить разрыв между [латентными]
способностями и [наблюдаемыми] результатами,
но и преодолеть его»²*

Для современной психометрики интерпретировать результаты работы с данными в терминах вероятности — привычный и устоявшийся подход. Он проявляется и в том, как описывается качество теста, и в формулировках обратной связи для участников тестирования. В психометрических исследованиях использование вероятностного подхода идет рука об руку с необходимостью учитывать ошибку измерения, т.е. разрыв между наблюдаемым поведением и той характеристикой, которую это наблюдаемое поведение должно было бы отражать. Не будет сильным упрощением сказать, что развитие психометрики — это осмысление ошибки измерения и усилия, направленные на ее сокращение.

С этой точки зрения возникновение и развитие байесовской психометрики оказывается ожидаемым, неизбежным и очень перспективным. Байесовское психометрическое моделирование объединяет принципы байесовской статистики с психометрическими моделями. Байесовская статистика уже активно

² "Much of this book is devoted to illuminating how adopting a probabilistic approach to psychometric modeling aids not only in representing the disconnect between [latent] proficiency and [observed] performance, but also in bridging that divide" (Levy R., Mislevy R.J. (2016) *Bayesian Psychometric Modeling*. Boca Raton, FL: Chapman and Hall/CRC. P. 9).

используется в машинном обучении и подарила нам возможность пользоваться «умными», т.е. самообучающимися, нейросетями. Использование байесовского подхода в психологии и образовании только набирает популярность [König, van de Schoot, 2018].

Байесовский подход часто рассматривается в противовес более привычному — частотному. Между ними существуют три базовых различия. Во-первых, в частотном подходе мы делаем вывод о вероятности наступления события только на основе собранных данных. Казалось бы, как может быть по-другому? Байесовский подход позволяет использовать не только уже существующие данные, но и наши предположения о том, как могут вести себя данные в будущем или что характерно для оцениваемого параметра в целом, вне зависимости от того, какая выборка нам попала. Иными словами, байесовский подход позволяет использовать априорную информацию.

Во-вторых, в частотном подходе мы оцениваем интересующий нас параметр как некое конкретное значение, возможно, с ошибкой оценки параметра (*confidence interval*). А в байесовском подходе интересующий исследователя параметр рассматривается как случайная величина, имеющая свое (апостериорное) распределение вероятности. В этом распределении можно выделить некое значение, например среднее, и определенный интервал для дальнейшей интерпретации (*credibility interval*).

В-третьих, существование распределения вероятности в байесовском подходе позволяет отойти от оценки статистической значимости и проверять гипотезу на основе апостериорного распределения вероятности.

Применение байесовской статистики распространялось в исследовательском поле по мере роста компьютерных мощностей: ограничениями в повсеместном использовании байесовского подхода выступают продолжительность расчетов и необходимость наличия больших вычислительных мощностей. В 1995 г. вышло первое издание книги Э. Гелмана и его коллег "*Bayesian Data Analysis*", а в 2013 г. опубликовано ее третье расширенное издание [Gelman et al., 2013]. В ней приведены концептуальные постулаты байесовского подхода, его математические обоснования, а также примеры использования в разных областях³. Однако Э. Гелман с коллегами не рассматривают вопросы, специфические для психометрических исследований.

В 2010 г. увидела свет книга Ж.-П. Фокса "*Bayesian Item Response Modeling. Statistics for Social and Behavioral Sciences*" [Fox, 2010]. В ней рассмотрены базовые и передовые психометри-

³ <http://www.stat.columbia.edu/~gelman/book/> Текст книги и сопроводительные материалы доступны по ссылке для некоммерческого использования.

ческие модели современной теории тестирования в преломлении к байесовской парадигме анализа данных⁴. Существуют материалы, специфичные и для других психометрических подходов, например для структурного моделирования⁵ [Muthén, Asparouhov, 2012].

В рецензируемой книге Р. Леви и Р. Мислеви *“Bayesian Psychometric Modeling”* впервые представлено систематическое описание байесовского подхода в психометрических исследованиях. Издание содержит основные положения байесовской статистики как на уровне концептуального осмысления вероятностного вывода, так и на уровне математических формул. При этом баланс между концептуальными рассуждениями, иллюстративными примерами и строгим статистическим описанием сохраняется для всех представленных методов психометрического анализа.

Книга состоит из двух частей: основные положения байесовского подхода (*Foundations*) и их применение в психометрическом моделировании на примере разных методов (*Psychometrics*).

Первая часть включает шесть крупных разделов. Первые три раздела посвящены введению в оценивание и байесовскую вероятность. В частности, описана методология доказательной аргументации в разработке тестов (*evidence-centered design*) и ее связь с вероятностным моделированием. Кроме того, рассматриваются особенности байесовского вывода, ключевые формы распределений, а также возможные проблемы и подходы к их решению, например при использовании априорных вероятностей.

Четвертый, пятый и шестой разделы первой части книги посвящены подготовке к работе с реальными данными. Так, рассматривается моделирование переменной с нормальным распределением и описывается использующийся при этом метод оценки параметров — алгоритм Монте-Карло по схеме марковской цепи (*Markov Chain Monte Carlo*, МСМС). Кроме этого, подробно представлено байесовское моделирование линейной регрессии с непрерывной зависимой переменной, которое служит базой для более сложных психометрических моделей.

Вторая часть книги посвящена разным психометрическим моделям в рамках байесовского подхода. Авторы начинают с описания байесовского моделирования классической теории тестирования (*classical test theory*), а затем переходят к представлению байесовского конфирматорного факторного анализа

⁴ <https://www.jean-paulfox.com/> сопроводительные материалы доступны по ссылке для некоммерческого использования.

⁵ <https://mc-stan.org/users/documentation/case-studies/sem.html>

(*confirmatory factor analysis*) и современной теории тестирования (*item response theory*), включая анализ заданий с дихотомическим и полиномическим начислением баллов, многомерные модели. Далее авторы рассматривают моделирование латентных классов (*latent class analysis*) с примером для дихотомических латентных и наблюдаемых переменных, а также довольно молодой метод анализа данных в психологии и образовании — моделирование байесовскими сетями (*Bayesian networks*).

Большое внимание в книге уделяется анализу качества модели. Этой теме посвящена отдельная глава (*Model Evaluation*), а также разделы (*Model-Data Fit*) в каждой главе, описывающей психометрические модели. Авторы рассмотрели и еще одну важную проблему психометрики — работу с пропущенными данными (*missing data modeling*). Байесовский подход служит опорой в ее решении не только в психометрике, но и при анализе данных в социальных науках в целом.

Сильная сторона книги — ее направленность на практическое применение. Авторы подробно описывают, как реализовать байесовские психометрические модели с помощью бесплатного программного пакета WinBUGS⁶. Этот пакет и схожая с ним система OpenBUGS предоставляют мощный аппарат для получения апостериорных вероятностей, однако при их использовании возможности последующего анализа данных ограничены. Более удобен пакет R2WinBUGS [Sturtz, Ligges, Gelman, 2005]: он позволяет применять WinBUGS через привычный интерфейс программы R⁷, а также использовать возможности R для последующего анализа данных.

Книгу Р. Леви и Р. Мислеви "*Bayesian Psychometric Modeling*" можно использовать как учебник: в конце каждой главы приведены несколько упражнений разного уровня трудности.

Одним из потенциальных недостатков книги выступает уровень ее математической строгости. Авторы предполагают, что читатель обладает базовыми знаниями в теории вероятности и статистике. Кроме того, некоторые из наиболее сложных тем могут быть трудными для читателей, не знакомых с байесовской статистикой.

В кратком отклике на рецензируемую книгу Д. Хатчисон [Hutchison, 2018] отмечает, что она подходит для тех, кто работает в частотном подходе и хотел бы узнать про байесовский, а также для студентов аспирантуры. Рекомендация в целом верная, но требует уточнения. Во-первых, профессионалам, не уверенным в надежности своего математического бэкграунда

⁶ Spiegelhalter D., Thomas A., Best N., Lunn D. (2003) WinBUGS User Manual.

⁷ R Core Team (2022) R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>

(а таких среди исследователей в психологии и образовании немало), для более легкого старта стоит начать с менее математически нагруженных источников, например [Schoot van de et al., 2014]. Во-вторых, тем, кто заинтересован в освоении новой для него психометрической модели, например латентного классового анализа, целесообразно сначала освоить частотный вариант оценки параметров и после этого перейти к байесовскому.

“*Bayesian Psychometric Modeling*” Р. Леви и Р. Мислеви — это отличный ресурс для всех, кто заинтересован в изучении байесовской статистики и ее применении в психометрике. Книга содержит исчерпывающее описание этого подхода и включает многочисленные примеры его практического применения. Она, возможно, окажется сложной для читателей без сильной математической подготовки, но погружение в нее принесет пользу каждому, кто заинтересован в передовом анализе данных в психологии и образовании.

Благодарности Спасибо сотрудникам Центра психометрики и измерений в образовании за замечательный подарок на день рождения, который позволил мне стать ближе к байесовскому подходу, — за бумажный экземпляр книги “*Bayesian Psychometric Modeling*”.

- References**
- Fox J.-P. (2010) *Bayesian Item Response Modeling. Theory and Applications*. New York, NY: Springer. https://doi.org/10.1007/978-1-4419-0742-4_5
- Gelman A., Carlin J.B., Stern H.S., Dunson D.B., Vehtari A., Rubin D.B. (2013) *Bayesian Data Analysis*. Boca Raton, FL: Chapman and Hall/CRC.
- Hutchison D. (2018) Bayesian Psychometric Modelling. *Journal of the Royal Statistical Society Series A*, vol. 181, no 2, pp. 550–550. <https://doi.org/10.1111/rssa.12344>
- König C., van de Schoot R. (2018) Bayesian Statistics in Educational Research: A Look at the Current State of Affairs. *Educational Review*, vol. 70, no 4, pp. 486–509. <http://dx.doi.org/10.1080/00131911.2017.1350636>
- Muthén B., Asparouhov T. (2012) Bayesian Structural Equation Modeling: A More Flexible Representation of Substantive Theory. *Psychological Methods*, vol. 17, no 3, pp. 313–335. <http://dx.doi.org/10.1037/a0026802>
- Schoot van de R., Kaplan D., Denissen J., Asendorpf J.B., Neyer F.J., van Aken M.A. (2014) A Gentle Introduction to Bayesian Analysis: Applications to Developmental Research. *Child Development*, vol. 85, no 3, pp. 842–860. <http://dx.doi.org/10.1111/cdev.12169>
- Sturtz S., Ligges U., Gelman A. (2005) R2WinBUGS: A Package for Running WinBUGS from R. *Journal of Statistical Software*, vol. 12, no 3, pp. 1–16. <http://dx.doi.org/10.18637/jss.v012.i03>