Item Response Theory

Course Syllabus

Instructor: Hau-Hung Yang

Course Information

Course Level:	Graduate
Credits:	3 credits
Prerequisites:	Graduate-level statistics (regression, maximum likelihood estimation); basic familiarity with R or Python
Software:	R (packages: `mirt`, `ltm`, `catR`, `TAM`)

Course Description

This graduate-level course provides a comprehensive introduction to item response theory (IRT), the dominant psychometric framework for modeling the relationship between latent traits and observed responses to test items. The course covers the theoretical foundations of major IRT models, parameter estimation methods, model fit evaluation, and practical applications including test construction, differential item functioning (DIF) analysis, and computerized adaptive testing (CAT). Emphasis is placed on understanding the mathematical underpinnings of IRT models, developing the ability to conduct IRT analyses with real data, and critically evaluating the assumptions and limitations of IRT in applied settings.

Course Objectives

Upon completion of this course, students will be able to:

1. Describe the key assumptions and mathematical formulations of major IRT models (1PL/Rasch, 2PL, 3PL, graded response, generalized partial credit).
2. Understand and apply parameter estimation methods, including marginal maximum likelihood (MML), conditional maximum likelihood (CML), and Bayesian estimation (EAP, MAP).
3. Evaluate model fit and conduct model comparison using appropriate indices and graphical diagnostics.
4. Perform DIF analysis to assess measurement fairness across groups.

5. Understand the principles of computerized adaptive testing and implement basic CAT simulations.
6. Apply IRT methods to real datasets using R and interpret results in substantive contexts.

Textbooks and References

- de Ayala, R. J. (2022). *The Theory and Practice of Item Response Theory* (2nd ed.). Guilford Press. **[Primary textbook]**
- Baker, F. B., & Kim, S.-H. (2017). *The Basics of Item Response Theory Using R*. Springer.
- van der Linden, W. J. (Ed.) (2016). *Handbook of Item Response Theory* (Vols. 1–3). CRC Press.
- Supplementary journal articles (distributed by topic).

Assessment

Homework assignments (theory + R exercises)	30%
Midterm examination	20%
Final examination	20%
Applied IRT project (written report + presentation)	20%
Paper reading summaries and class discussion	10%

Weekly Schedule

Week	Topic
1	Introduction: From classical test theory to IRT; limitations of CTT; overview of latent trait models
2	The Rasch model (1PL): Model formulation, sufficient statistics, specific objectivity, and the conditional maximum likelihood estimator
3	The two-parameter logistic (2PL) model: Item discrimination and difficulty; item characteristic curves (ICCs) and item information functions
4	The three-parameter logistic (3PL) model: The guessing parameter; practical considerations and identification issues
5	Parameter estimation I: Joint, conditional, and marginal maximum likelihood estimation; the EM algorithm
6	Parameter estimation II: Bayesian estimation methods (EAP, MAP); prior specification and its impact
7	Ability estimation: Maximum likelihood, Bayesian, and weighted likelihood estimators; standard errors and confidence intervals
8	Test and item information: Fisher information, test information function, and its role in test design
9	*Midterm Examination*
10	Model fit assessment: Likelihood-ratio tests, information criteria (AIC, BIC), item fit statistics, and graphical diagnostics
11	Models for polytomous data I: The graded response model (GRM)
12	Models for polytomous data II: The (generalized) partial credit model; the nominal response model
13	Differential item functioning (DIF): Concepts, IRT-based detection methods (Lord's χ^2, likelihood-ratio test, Raju's area measure)
14	Test construction and item bank development: Item selection criteria, information-based design, and equating
15	Computerized adaptive testing I: Principles, item selection algorithms (maximum information, Bayesian), stopping rules
16	Computerized adaptive testing II: Practical issues—item exposure control, content balancing; CAT simulation in R
17	Advanced topics and current developments: Multidimensional IRT; connections to factor analysis; online calibration; recent applications of machine learning in IRT
18	*Final Examination*

Applied IRT Project

Each student (or group of two) will complete an applied IRT project involving a real or publicly available dataset. The project report (approximately 10–15 pages) should include:

1. A description of the dataset, research context, and the purpose of the analysis.
2. Preliminary item analysis using classical test theory indices.
3. Model fitting: Fit at least two IRT models (e.g., Rasch vs. 2PL, or GRM vs. GPCM) and compare them using appropriate fit indices.
4. Parameter estimation results with interpretation of item and ability parameters.
5. At least one applied analysis, such as: DIF detection, test information evaluation, or a CAT simulation based on the fitted item bank.
6. A discussion of the findings, limitations, and implications for test development or revision.

Students will present their projects during Week 17. All analyses must be documented in reproducible R code.

Course Policies

- All homework and project analyses must be submitted as reproducible R scripts or R Markdown documents.
- Late submissions will incur a 10% penalty per day, up to a maximum of three days.
- Students are expected to read assigned papers before class and participate actively in discussions.
- AI tools may be used for learning and coding assistance, but students must demonstrate independent understanding of the methods and cite any AI-assisted work.