EPSY 551: Item Response Theory / Rasch Measurement

4 Credit Hours. CRN: 43708.

Semester: Fall 2021	**Professor: George Karabatsos (home page)**
Time when weekly new video lecture material is presented: every Tuesday 5:00-8:00pm	**Phone: 312-413-1816**
Classroom: Online asynchronous.	**E-mail: georgek@uic.edu**
Computer Lab: ETMSW2027.	**Office Hours: Tuesday 2-5pm (EPASW 1034)**

Course Description:
Item Response Theory (IRT) models continue to see vastly many applications in the educational, psychological, health, and other social sciences. This course teaches how to apply classical and contemporary IRT models for the psychometric analysis of data arising from examinees' responses to items of a test (e.g., examination, rating scale questionnaire), and for the analysis of judge ratings of examinee performance on test items. In an IRT model analysis of item response data, the main objective is to estimate each examinee's ability of the test, each item's difficulty level, and perhaps, estimate the regression coefficients of one or more (examinee-level, item-level, and/or judge/rater level) predictor variables. IRT models that are covered by this course include Rasch models, the 2-parameter logistic model, the 3-parameter logistic model, the nonparametric monotone homogeneity model (via kernel regression), graded response models, the generalized partial credit model, Rasch models with predictor variables, as well as finite-mixture and infinite-mixture (e.g., Dirichlet process mixture) versions of these IRT models. While this course focuses and grades on applications of various IRT models, this course also covers the theoretical foundations of IRT models, and the modern methods used to estimate the parameters of IRT models from data (including point-estimation and Bayes posterior estimation). These models will be illustrated through the analysis of data using various software packages including PARSCALE, WINSTEPS, eRm, FACETS, mixRasch, HLM, IRTPRO, Bayesian Regression, and other software packages that will be made available to students free-of-charge. The course does not assign special preference to certain IRT or Rasch models, but instead allows students to exercise their respective individual preferences. Assignment (exam) problems will be practiced and worked on in each class lecture/session.

Prerequisites: At least two graduate courses on quantitative methods.

Books and suggested readings. (The book below, in bold, is available in the UIC bookstore):
De Boeck, P., and Wilson, M. (2004). Explanatory Item Response Models. Boca Raton, FL: Chapman and Hall/CRC.
Embretson, S.E., & Riese, S.P. (2000). *Item Response Theory for Psychologists*. Lawrence Erlbaum Associates, Mahwah, NJ.
Karabatsos, G., Walker, S.G. (2009). A Bayesian Nonparametric Approach to Test Equating. *Psychometrika*, *74*, 211-232.
Levy, R., and Mislevy, R.J. (2016). *Bayesian Psychometric Modeling*. Boca Raton, FL: Chapman and Hall/CRC.
Meijer, R R., & Nering, M.L. (1999). Computerized adaptive testing: Overview and introduction. *Applied Psychological Measurement*, *23*, 187-194.
van der Linden, W.J. (2016). *Handbook of Item Response Theory, Volume One*: *Models*. Boca Raton, FL: Chapman and Hall/CRC.
van der Linden, W.J. (2016). *Handbook of Item Response Theory, Volume Two*: *Statistical Tools*. Boca Raton, FL: Chapman and Hall/CRC.
van der Linden, W.J. (2018). *Handbook of Item Response Theory, Volume Three: Applications*. Boca Raton, FL: Chapman and Hall/CRC.

Important Dates:
August 24, T, Instruction begins (lecture material is first presented on Tuesday evening, 5-8pm).
October 29, F, Last day to drop courses and receive a W on record
November 25–26, Th–F, Thanksgiving holiday. No classes.
December 3, F Instruction ends.
December 6-10, Final Exam Week: Final Exam due by Wednesday 11:59pm.

COURSE SCHEDULE

Week	Topic
1	**What is an IRT model? IRT foundations** Basic data structure in IRT modeling. The general unidimensional IRT model. The item response function (IRF) (item characteristic curve, item-step response function (ISRF), item category-response function.
2	**IRT foundations (continued)** -- The three properties of all psychometric models (unidimensionality, local independence, monotonicity of the IRF/ISRF). -- Invariant item ordering. -- Examples of IRT models under this general framework. **IRT/Rasch models for data analysis** Parameter maximum likelihood estimation methods for IRT models. Model selection methods

3	**IRT/Rasch models for dichotomous item scores** Rasch, 2-parameter, and 3-parameter logistic models. Illustrative applications of models on real data.
4	**IRT/Rasch methods for polytomous item scores** Real data illustrations of various IRT models for polytomous item scores. Graded response models, generalized partial credit model, Rasch rating scale model. **EXAM #1 is due in Blackboard (on dichotomous parametric IRT models).**
5	**Nonparametric IRT: Kernel regression approach** Kernel estimation of the IRF, the ISRF, and the category response function. Investigating measurement unidimensionality (investigating IRF/ISRF monotonicity). Estimating examinee ability, and item difficulty. Illustrative applications of models on real data.
6	**Nonparametric IRT: Kernel regression approach** Illustrative applications of models on real data.
7	**IRT/Rasch modeling with covariates via Hierarchical Linear Models (HLM)** Rasch model as a Hierarchical Linear Model. (Rasch) analysis of test items, rating scales, and judge ratings. Investigating Item Bias (Differential Item Functioning). Comparing test performance across different groups of respondents. Incorporating additional predictor variables in psychometric analysis. Illustrative applications of models on real data.
8	**IRT/Rasch modeling with predictor variables via HLM** Illustrative applications of models on real data. **EXAM #2 is due in Blackboard (on polytomous, kernel nonparametric, or hierarchical IRT models).**
9	**Finite Mixture (latent class) IRT models** Examples of models. Illustrative applications of models on real data.
10	**Bayesian nonparametric (infinite-mixture) Rasch models for data analysis** Bayesian estimation of the posterior distribution. Various applications of Bayesian semiparametric mixed Rasch models. Analysis of test items, rating scales, and judge ratings. Investigating Item Bias (Differential Item Functioning). Comparing test performance across different groups of respondents. Incorporating additional predictor variables in psychometric analysis.
11	**Bayesian nonparametric mixed Rasch models for data analysis** Illustrative applications of models on real data.
12	**Equating Test Scores and test items** The equipercentile approach to test equating (with bootstrap confidence intervals). Linear equating. Rasch item equating. Bayesian nonparametric equating.
13	**Computer Adaptive Testing using IRT** **EXAM #3 is due in Blackboard (on mixture IRT models).**
14	**Student Presentations**
15	**Student Presentations**
16	***FINAL PAPER DUE (Exam week).*** Please submit final paper in Blackboard.

Course Assignments

Exams 1, 2, and 3: Each take-home exam gives you the opportunity to practice with applying and explaining various IRT/Rasch models through analysis of real data sets.
Short paper and presentation: A 7-page (double-spaced) paper, on applying and explaining various IRT/Rasch models, to be presented in class near the end of the semester.
For each exam, and for the the paper, please show all work and relevant output, and place all raw output in the Appendix.
Note: You have the option to not do an in-class presentation, in which case you would be required to submit a 12-page (double-spaced) paper.
That being said, historically, one of the more enjoyable aspects of the course has been watching and giving the in-class presentations on IRT models.

Grading Policy: Exams 1, 2, 3, and the final paper (with presentation) are each worth 22.5% of the final grade (together, they are worth 90%).
Class participation is worth the remaining 10% of the final grade. Final grades will be given out according to the following scale:

A	90% - 100%
B	79% - 89%

C	68% - 78%
D	57% - 67%
F	56% - Lower

The amount of student class participation will be used to decide borderline grades.
Students will spend substantial amounts of time reading, and on the computer.
It is assumed that students will exert individual initiative in solving computing/analysis problems as they arise.
There are no exceptions to the above grading scale, and no extra credit work will be accepted.
Incompletes will be considered for students with extenuating circumstances.
Poor performance on assignments will not be considered in a request for an incomplete.

*The instructor reserves the right to make any changes in the course he determines academically advisable.
Changes will be announced in class. It is your responsibility to keep up with any changed policies.

DATA ANALYSIS PAPER/PRESENTATION:
The data analyses will consist of the relevant output from the software programs and a complete report stating the results.
You may supply your own data or you may solicit faculty (education or other) for data.
Some data sources: Open Psychometrics, PIRLS, TIMSS, PISA.
The final paper should be at least 7 double-spaced pages (not including Appendix), using 1-inch margins and APA or other standard scientific style.
Please discuss only the relevant results of your analysis, within the main body of the paper.
Please put all data analysis output in the Appendix of your paper.
The presentation has a limit of 20 minutes (about 15 PowerPoint slides).
Please hand me a hard-copy of your PowerPoint presentation on the day of your presentation.

The **paper/presentation** must at least include the following sections:

Introduction
Describe in detail the substantive problems you will be addressing in this research study (5 points).

Methods
Describe sample characteristics (3 points).
Describe the items on your test(s) (including their number and scoring format) (3 points).
Describe the construct you intend to measure with the test(s) (3 points).
For data analysis, use one or more IRT models for data analysis.
Fully describe your IRT model(s) and the parameter estimation methods (15 points)
If you intend to equate test scores, fully describe the equating methods you will implement.
If you will implement one or more IRT models and plan to compare them, describe the model selection method you will use.

Results
Summarize the results of your analysis, including estimates of IRT model parameters (10 points).
If necessary, justify any modifications you make to your test (e.g., removing items).
Please discuss only the relevant results of your analysis. (15 points)

Discussion
What modifications (if any) would improve the test? (3 points)
What are the implications of your study, with respect to the measurement and applications in the field of interest? (3 points)

Please provide appropriate handouts and develop meaningful overheads for your presentation.

Disability Services:
UIC strives to ensure the accessibility of programs, classes, and services to students with disabilities.
Reasonable accommodations can be arranged for students with various types of disabilities,
such as documented learning disabilities, vision, or hearing impairments, and emotional or physical disabilities.
If you need accommodations for this class, please let your instructor know your needs and s/he
will help you obtain the assistance you need in conjunction with the Office of Disability Services (1190 SSB, 413-2183).