APSTA-GE-2094
Modern Approaches in Measurement
APSY-GE 2025
Psychological Measurement

Klint Kanopka, Ph.D. (he/him)
Assistant Professor of Applied Statistics

Spring 2026

1 Course Description

The course provides students with the software skills and theoretical knowledge required to apply latent construct measurement techniques in R. The course assumes a working knowledge of R, linear regression, and basic probability. Continuous and categorical latent variables, dimensionality reduction, clustering, finite mixture, and diagnostic classification models are covered. The course focuses on cross-sectional applications of measurement techniques, with examples drawn from education, psychology, and social science.

2 Student Learning Outcomes

During the course of this class:

1. Students will identify and define continuous and categorical latent constructs.
2. Students will relate unsupervised learning techniques and measurement models.
3. Students will compare and contrast competing measurement models.
4. Students will select, apply, and evaluate measurement models to analyze real data.

3 Course Logistics

3.1 Teaching Team

- **Klint Kanopka** - Instructor
  - klint.kanopka@nyu.edu
  - Office Hours: W 2p-3p @ Kimball, Rm 205W

3.2 Meeting Times and Locations

In-person lecture attendance is required, as lectures are not recorded by default. To request a lecture be recorded, please contact me directly.

- **Lecture:** Th 9.15a-12.15p @ Silver, Rm 411

3.3 Required and Recommended Reading

- *A Short Guide to Item Response Theory Models*[28] is a new book covering a great deal of introductory item response theory and modeling. https://link.springer.com/book/10.1007/978-3-031-87271-6[1]
- *Elements of Statistical Learning* [18] covers the unsupervised machine learning core and estimation. https://hastie.su.domains/ElemStatLearn/
- *Handbook of Modern Item Response Theory* [31] covers a wider variety of item response models. https://link.springer.com/book/10.1007/978-1-4757-2691-6[2]

3.4 Prerequisites

Students should be experienced with the `R` Programming Language [24], probability, regression, and data visualization. Additional experience in survey methodology and unsupervised machine learning is helpful but not required.

4 Grading

This is a three-credit course. Table 1 contains the proportion of the final grade coming from each assignment category. Table 2 contains the transformation from the weighted proportion of earned points to letter grades.

[1] You may need to be on NYU servers or using a VPN
[2] You may need to be on NYU servers or using a VPN

Category	p
Problem Sets	0.9
Participation (Lecture)	0.1

Table 1: Final grade weighting scheme

	G^-	G	G^+
A	$[.895, .945)$	$[.945, 1]$	—
B	$[.795, .825)$	$[.825, .865)$	$[.865, .895)$
C	$[.695, .725)$	$[.725, .765)$	$[.765, .795)$
D	$[.600, .640)$	$[.640, .670)$	$[.670, .695)$
F	—	$[0, .600)$	—

Table 2: Grading Scale

4.1 Problem Sets

The course contains eight equally-weighted problem sets (PS0-PS7). PS0 is a one-week assignment. PS2-PS6 are two-week assignments. PS7 is longer, due during finals week, and closer to a "project."

4.1.1 Submission Guidelines

All assignments are to be submitted on Gradescope as a `.pdf` and adhere to the following conventions:

1. **Submit assignments by the deadline.** Assignment deadlines are typically Fridays at 11.59p. Late assignments will receive a 10% penalty per day up to a maximum of three days.

2. **Name your files correctly.** For each homework assignment, name your files using the convention: `LastName_FirstInitial_PS#.ext`

3. **Put your name and the assignment in the text of the file.** These should occupy the author and title fields, respectively.

4. **Submit a compiled file on Gradescope.** Source files are Quarto documents and have the extension `.qmd`. Please retain these to submit upon request. Submitted files should be compiled to `.pdf` and generated from the `.qmd` template provided. As an example, for the first homework, I would create two files:

   - `Kanopka_K_PS0.pdf`
   - `Kanopka_K_PS0.qmd`

   I would submit `Kanopka_K_PS0.pdf` on Gradescope and retain a copy of `Kanopka_K_PS0.qmd` locally.

5. **If requested, your source file should run without modification.** If you load a file, please use relative paths. Do not load libraries unless explicitly requested in an assignment.

6. **Do not use `install.packages()` calls in your assignments.** Please install packages locally and then load them in your assignment source code.

4.2 Extra Credit

This course allows extra credit to be earned in any of the following three ways:[3]

1. The first person to report typographical errors or mistakes in any of my course materials will receive one (1) extra credit point per error (I expect there to be a lot of them).

2. "Valuable suggestions" will also result in one (1) extra credit point and acknowledgment in a footnote.

3. Reporting a coding error with a solution or a valuable suggestion that requires additional labor (code snippets, a figure, a dataset, etc.) will be worth more, up to a maximum of 2^8 points. This amount will be scaled based on the nature and duration of required labor.

These should be claimed in the `#2094-measurement-errata` Slack channel.

5 Course Policies

5.1 Academic Integrity

I take academic integrity incredibly seriously. Please review NYU Steinhardt's Academic Policies and Procedures for more information on specific policies, the disciplinary process, and sanctions.

5.2 Collaboration

I **strongly** encourage students to form study groups. Students may discuss and work on problem sets in groups. Each student must (1) report at the top of each question what other students they consulted with, (2) write their code and solution independently, and (3) understand their work well enough in order to reconstruct it entirely on their own.

[3] This policy is inspired by the Knuth Reward Check

5.3 AI Tool Policy

All assignments should be your own original work, but students should feel comfortable consulting AI tools to support their work. The expectation is that students will write their own software implementations, interpret their results, and develop their own narrative responses, but may use AI tools to help with tasks like debugging and refining their grammar. If you have a question about a specific tool or use case, please reach out to me.

5.4 Students with Disabilities

Students with physical or learning disabilities are required to register with the Moses Center for Student Accessibility, 726 Broadway, 2nd Floor, (212-998-4980 and online at `http://www.nyu.edu/csd`). They must present a letter from the Center to the instructor at the start of the semester to be considered for appropriate accommodation.

5.5 Mental Health Statement

If you are experiencing undue personal and/or academic stress during the semester that may be interfering with your ability to perform academically, the NYU Wellness Exchange (212-443-9999) offers a range of services to assist and support you. I am available to speak with you about stresses related to your work in my course, and I can assist you in connecting with the Wellness Exchange. Additionally, if you anticipate any challenges with completing the assignments, readings, exams and other work required in this course, I encourage you to register with the Moses Center (212-998-4980) in advance so that you may be granted the proper academic accommodations.

5.6 Inclusion

NYU values an inclusive and equitable environment for all our students. I hope to foster a sense of community in this class and consider it a place where individuals of all backgrounds, beliefs, ethnicities, national origins, gender identities, sexual orientations, religious and political affiliations, and abilities will be treated with respect. I intend that all students' learning needs be addressed in and out of class and that the diversity students bring to this class be viewed as a resource and strength. Please feel free to speak with me if this standard is not being upheld.

6 Course Calendar

Table 3 contains topics, assignments, and deadlines. Note that this is subject to change.

Week	Lecture	Assignments
1	Topics: • What is measurement? • What is psychometrics? • Reliability • Latent variables Reading: • *A Short Guide to Item Response Theory Models*[28] – Chapter 7: Classical Test Theory • *Construct Validity in Psychological Tests*[7] • *An Argument-based Approach to Validity*[19] • *The Concept of Validity*[6]	Released: • HW0
2	Topics: • Classical Test Theory (CTT) • Principal Component Analysis (PCA) • Factor Analysis (EFA) Reading: • *Classical test theory in historical perspective*[27] • *An Introduction to Statistical Learning*[17] – Chapter 12, Introduction – Section 12.1 – Section 12.2 • *The Elements of Statistical Learning*[18] – Chapter 14, Introduction – Section 14.7, Introduction and Factor Analysis – Section 14.5	Released: • HW1 Due: • HW0

3	Topics: • Rasch Measurement • Item Response Theory • The EM Algorithm Reading: • *A Short Guide to Item Response Theory Models*[28] – Chapter 1: Introduction – Chapter 2: The Binary Rasch Model – Chapter 3: Extensions of the Rasch Model and Alternative Binary Models • *Handbook of Modern Item Response Theory*[31] – Chapter 1 • *A Rasch Primer: The Measurement Theory of Georg Rasch*[22] – Part I and Part II—Stop when you get to Polytomous Rasch models	
4	Topics: • Two and Three Parameter Item Response Models • Multidimensional Item Response Theory • Explanatory Item Response Models Reading: • *A Short Guide to Item Response Theory Models*[28] – Chapter 3: Extensions of the Rasch Model and Alternative Binary Models – Chapter 11: Explanatory Item Response Models • *The Past and Future of Multidimensional Item Response Theory*[25] • *Explanatory Item Response Models: A Brief Introduction*[34]	Released: • HW2 Due: • HW1

5	Topics: • Polytomous Item Response Models Reading: • *A Short Guide to Item Response Theory Models*[28] – Chapter 4: Ordinal Models – Chapter 5: Extended Ordinal Models – Chapter 6: The Thresholds Model • *An NCME Instructional Module on Polytomous Item Response Theory Models*[23] • *The Estimation of Item Response Models with the* `lmer` *Function from the* `lme4` *Package in* `R`[9]	
6	Topics: • Getting creative about "item responses" • IRTrees • Ideal Point Models • Elo Systems • Paired Comparisons Reading: • *A Short Guide to Item Response Theory Models*[28] – Chapter 9: Tree-based Item Response Models • *IRTrees: Tree-based Item Response Models of the GLMM Family*[8] • *Elo ratings and the sports model: A neglected topic in applied probability?*[4] • *Ideal Points and American Political Development: Beyond DW-NOMINATE*[1]	Released: • HW3 Due: • HW2

7	Topics: • Measurement of text • Measurement with text Reading: • *Essay content and style are strongly related to household income and SAT scores: Evidence from 60,000 undergraduate applications*[2] • *Text as Data Methods for Education Research*[11] • *Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts*[13] • `stm`: *An* `R` *Package for Structural Topic Models*[26] • *Word Embeddings Quantify 100 Years of Gender and Ethnic Stereotypes*[12]	
8	Topics: • Categorical Latent Variables • Cluster Analysis • k-Means, Agglomerative, and Divisive Clustering Reading: • *An Introduction to Statistical Learning*[17] – Chapter 12.4: Clustering Methods – Chapter 12.5: Lab Unsupervised Learning • *Sociological Classification and Cluster Analysis*[3] • *The Occupational Composition of American Classes: Results from Cluster Analysis*[32]	Released: • HW4 Due: • HW3

9	Topics: • "Soft" and Probabilistic Clustering • Mixture Models Reading: • *Hands on Machine Learning with* R[5] – Chapter 22: Model-based Clustering – Chapter 20: k-Means Clustering – Chapter 21: Hierarchical Clustering	
10	Topics: • Latent Class Analysis Reading: • *Latent Class Analysis: A Guide to Best Practice*[33]	Released: • HW5 Due: • HW4
11	Topics: • Categorical Latent Traits • Cognitive Diagnostic Models Reading: • *An Application of Latent Class Models to Assessment Data*[14] • *Continuous and Discrete Latent Structure Models for Item Response Data*[15] • *DINA Model and Parameter Estimation: A Didactic*[10] • GDINA*: An* R *Package for Cognitive Diagnosis Modeling*[21]	

12	Topics: • Process Data • Response Time Modeling Reading: • *Implications of considering Response Process Data for Greater and Lesser Psychometrics*[20] • *A Data Science Perspective on Computational Psychometrics*[16] • *Conceptual Issues in Response-Time Modeling*[30] • *A Lognormal Model for Response Times on Test Items*[29]	Released: • HW6 Due: • HW5
13	Topics: • Special Topics Reading: • TBD	
14	Topics: • Special Topics Reading: • TBD	Due: • HW6

Table 3: Course Calendar

References

[1] D. Aldous. Elo ratings and the sports model: A neglected topic in applied probability? 2017.

[2] A. Alvero, S. Giebel, B. Gebre-Medhin, A. L. Antonio, M. L. Stevens, and B. W. Domingue. Essay content and style are strongly related to household income and sat scores: Evidence from 60,000 undergraduate applications. *Science advances*, 7(42):eabi9031, 2021.

[3] K. D. Bailey. Sociological classification and cluster analysis. *Quality & Quantity*, 17(4), 1983.

[4] D. A. Bateman and J. Lapinski. Ideal points and american political development: Beyond dw-nominate. *Studies in American Political Development*, 30(2):147–171, 2016.

[5] B. Boehmke. Hands-on machine learning with r, 2019.

[6] D. Borsboom, G. J. Mellenbergh, and J. Van Heerden. The concept of validity. *Psychological review*, 111(4):1061, 2004.

[7] L. J. Cronbach and P. E. Meehl. Construct validity in psychological tests. *Psychological bulletin*, 52(4):281, 1955.

[8] P. De Boeck and I. Partchev. Irtrees: Tree-based item response models of the glmm family. *Journal of Statistical Software*, 48:1–28, 2012.

[9] P. De Boeck, M. Bakker, R. Zwitser, M. Nivard, A. Hofman, F. Tuerlinckx, and I. Partchev. The estimation of item response models with the lmer function from the lme4 package in r. *Journal of Statistical Software*, 39: 1–28, 2011.

[10] J. De La Torre. Dina model and parameter estimation: A didactic. *Journal of educational and behavioral statistics*, 34(1):115–130, 2009.

[11] L. Fesler, T. Dee, R. Baker, and B. Evans. Text as data methods for education research. *Journal of Research on Educational Effectiveness*, 12 (4):707–727, 2019.

[12] N. Garg, L. Schiebinger, D. Jurafsky, and J. Zou. Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences*, 115(16):E3635–E3644, 2018.

[13] J. Grimmer and B. M. Stewart. Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political analysis*, 21(3):267–297, 2013.

[14] E. Haertel. An application of latent class models to assessment data. *Applied Psychological Measurement*, 8(3):333–346, 1984.

[15] E. H. Haertel. Continuous and discrete latent structure models for item response data. *Psychometrika*, 55(3):477–494, 1990.

[16] J. Hao and R. J. Mislevy. A data science perspective on computational psychometrics. In *Computational Psychometrics: New Methodologies for a New Generation of Digital Learning and Assessment: With Examples in R and Python*, pages 133–158. Springer, 2022.

[17] T. Hastie, R. Tibshirani, and J. Friedman. An introduction to statistical learning. 2009.

[18] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman. *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer, 2009.

[19] M. T. Kane. An argument-based approach to validity. *Psychological bulletin*, 112(3):527, 1992.

[20] R. Levy. Implications of considering response process data for greater and lesser psychometrics. *Educational Assessment*, 25(3):218–235, 2020.

[21] W. Ma and J. de la Torre. Gdina: An r package for cognitive diagnosis modeling. *Journal of Statistical Software*, 93:1–26, 2020.

[22] R. Mead. A rasch primer: The measurement theory of georg rasch. *Psychometrics services research memorandum*, 1, 2008.

[23] R. D. Penfield. An ncme instructional module on polytomous item response theory models. *Educational Measurement: Issues and Practice*, 33(1):36–48, 2014.

[24] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2018. URL `https://www.R-project.org/`.

[25] M. D. Reckase. The past and future of multidimensional item response theory. *Applied Psychological Measurement*, 21(1):25–36, 1997.

[26] M. E. Roberts, B. M. Stewart, and D. Tingley. Stm: An r package for structural topic models. *Journal of statistical software*, 91:1–40, 2019.

[27] R. E. Traub. Classical test theory in historical perspective. *Educational Measurement*, 16:8–13, 1997.

[28] G. Tutz. *A short guide to item response theory models*. Springer Nature, 2025.

[29] W. J. Van der Linden. A lognormal model for response times on test items. *Journal of Educational and Behavioral Statistics*, 31(2):181–204, 2006.

[30] W. J. Van Der Linden. Conceptual issues in response-time modeling. *Journal of Educational Measurement*, 46(3):247–272, 2009.

[31] W. J. van der Linden and R. K. Hambleton. *Handbook of modern item response theory*. Springer Science & Business Media, 2013.

[32] R. Vanneman. The occupational composition of american classes: Results from cluster analysis. *American Journal of Sociology*, 82(4):783–807, 1977.

[33] B. E. Weller, N. K. Bowen, and S. J. Faubert. Latent class analysis: a guide to best practice. *Journal of black psychology*, 46(4):287–311, 2020.

[34] M. Wilson, P. De Boeck, and C. H. Carstensen. Explanatory item response models: A brief introduction. *Assessment of competencies in educational contexts*, pages 91–120, 2008.