ST 790, Introduction to Dynamic Treatment Regimes
Fall 2020

Course:
Lectures: T-Th, 3:00 pm - 4:15 pm
Website: https://dtrcourse.wordpress.ncsu.edu/
Prerequisites: ST 702, ST 705 (or PhD level background in probability theory and statistical inference), basic knowledge of statistical models, programming skills

Instructor: Marie Davidian, davidian@ncsu.edu, www4.stat.ncsu.edu/~davidian
Online office hours Th 1:00 pm - 2:00 pm and by appointment

Teaching Assistant: Peter Norwood, pnorwoo@ncsu.edu,
Online office hours T 4:30 pm - 5:30 pm

Goal: This course will provide a comprehensive introduction to methodology for data-based development and evaluation of dynamic treatment regimes. A dynamic treatment regime is a set of sequential decision rules, each corresponding to a key point in a disease or disorder process at which a decision on the next treatment action must be made. Each rule takes patient information to that point as input and returns the treatment s/he should receive from among the available options, thus tailoring treatment decisions to a patient's individual characteristics. Dynamic treatment regimes formalize how clinicians make decisions in practice by synthesizing evolving information on a patient and are thus of considerable importance in precision medicine. Dynamic treatment regimes are also relevant in other contexts in which sequential decisions on interventions or policies must be made, as in education, engineering, economics and finance, and resource management.

Methods for estimation of dynamic treatment regimes and in particular optimal treatment regimes from data will be motivated and developed through a formal time-dependent causal inference framework. The gold standard study design for developing and evaluating regimes is the sequential multiple assignment randomized trial (SMART), considerations for which will be discussed. Inference for optimal treatment regimes is a nonstandard statistical problem and is thus notoriously difficult; an introduction to this challenge will be presented. Examples throughout the course will be drawn from cancer and other chronic disease research and research in the behavioral, educational, and other sciences.

Use of the comprehensive R package DynTxRegime to implement many of the methods discussed in the lectures will be introduced.

Students completing this course will have a foundation in causal inference and fundamental results and methods for dynamic treatment regimes that will provide the basis for study of the rapidly evolving literature on dynamic treatment regimes and precision medicine.

Text: Lecture notes prepared by the instructor. These will be available on the course website. The notes are based on the book Dynamic Treatment Regimes: Statistical Methods for Precision Medicine (2020) by Tsiatis, Davidian, Holloway, and Laber, which will serve as an optional reference.

Course delivery: This course will be delivered online in synchronous mode: class meetings will take place on Zoom on the scheduled days/time above, and students are expected to attend lectures remotely. Lectures will be recorded and made available on the course website shortly after each class meeting; access to recorded lectures will be password protected and limited only to registered students. The recordings are meant primarily as a resource for students to revisit the material rather
than as an alternative to attendance in real time. The recordings will also allow students who must miss a lecture due to illness or other conflict to review the material they missed.

Please note that recordings will capture the entire lecture and will include questions and comments of both instructor and students. Recorded lectures will be accessible until the end of the fall semester.

**Communication:** Office hours for both the instructor and TA will be held online on Zoom. Students can email the instructor anytime to make an appointment to meet outside of office hours on Zoom or to ask questions about course material or homework assignments.

Other than Zoom, the primary mode of communication between the instructor and students will be email. Students should check their email often for announcements from the instructor, and student should feel free to email the instructors with questions or concerns.

**Grading:** Final grade will be determined by the Final Score = 0.20 \times A + 0.50 \times P + 0.25 \times F + 0.05 \times D, where A is the average on analytical homework problems, P is the average on data analysis homework problems, F is the score on the final project, and D is instructor discretion score, where each is scored out of 100. The instructor's discretion portion will be based on attendance, participation in class, and instructor's assessment of mastery of the material. There will no midterm or final exam. Conversion of this score into a letter grade will be made according to the following tentative grading scale (the upper score in each range except A+ belongs to the next highest grade): A+, 100; A, 92-99; A-, 90-92; B+, 88-90; B 82-88; B-, 78-82. Scores below 70 will be handled on a case-by-case basis.

Please note that the plan for the course and grading is subject to change if circumstances related to the pandemic dictate.

Chronic absenteeism will result in at least a 5 point reduction in the Final Score, as determined by the instructor. If you you must miss a lecture due to illness, job interview, etc, the instructor would appreciate being informed in advance of the lecture if possible.

Auditors must attend class and attempt and turn in analytical problems on the homeworks. Auditors are not required to complete data analysis problems or the final project.

**Homework:** There will be four (4) homework assignments. Each homework assignment will involve an analytical problems section (denoted as A in the final score above) and a data analysis problems section (denoted as P in the final score above). The analytical problems section will comprise one or more problems involving derivations, proofs, and simulation studies. The data analysis problems section will involve carrying out analyses of data using the methods discussed in the lectures as programmed by you or implemented in the R package **DynTxRegime**. For problems involving programming, both the program and its output should be turned in, along with interpretation of the results as dictated by the problem. Unexcused late homework will be discounted by 50%.

Students are permitted and even encouraged to work together on homework; however, each student must prepare his/her own solutions. Blind copying of the work of other students demonstrates that the student doing the copying is not serious about developing the independence required for a PhD and has obvious disadvantages for the final project, not to mention mastery of the material.

The analytical problems section and data analysis problem section should be prepared as separate pdf files. Ideally, homework assignments should be typed (e.g., using **LaTeX**). Completed homeworks should be emailed to the instructor **at or prior to the beginning of class** on the due date.

Tentative assignments/due dates are as follows; definitive information will be posted on the course website.
Final Project (tentative): The final project will likely involve groups of students reading and presenting current research in the literature during the last week of the semester.

Computing/Software: Students are free to program in the language of their choice. However, prior familiarity with R is desirable, as we will use the R package DynTxRegime extensively.

Tentative Schedule: August 11 - November 12, 2020

08/11 - 08/13 – 1. Introduction
- Motivation (precision medicine, clinical decision making)
- Meaning of “dynamic” (dynamic vs static regimes)
- Basic framework ($\kappa$-decision regime, notation, optimal regime)
- Data (SMARTs, observational studies)

08/13 - 08/20 – 2. Preliminaries: Basic Causal Inference
- Statistical models
- Point exposure studies
- Potential outcomes and causal inference (causal effect, average causal effect, identifiability assumptions)
- Estimation of causal effects via outcome regression
- Review of M-estimation
- Estimation of causal effects via the propensity score (stratification, inverse probability weighted estimator)
- Doubly robust estimation of causal effects (augmented inverse probability weighted estimators, efficient AIPW estimator)

- Treatment regimes for a single decision point (potential outcomes, value)
- Estimation of the value of a fixed regime (identifiability assumptions, outcome regression estimator, IPW/AIPW estimators)
- Characterization of an optimal regime (in terms of potential outcomes, observed data)
- Estimation of an optimal regime (regression, A-learning, direct search IPW/AIPW, nonregularity)
- Use of the R package DynTxRegime

09/03 - 09/10 – 4. Single Decision Treatment Regimes: Additional Methods
- Optimal regimes from a classification perspective (IPW/AIPW estimators)
- Outcome weighted learning
- Interpretable regimes via decision lists
- Additional approaches
- Use of the R package DynTxRegime

09/10 - 09/22 – 5. Multiple Decision Treatment Regimes: Fundamentals
- Treatment regimes for multiple decision points (definition, recursive representation)
- Statistical framework (potential outcomes, feasible set of treatments and classes of regimes, value, data and data sources, identifiability assumptions)
- The g-computation algorithm
- Estimation of of the value of a fixed regime (via g-computation, IPW/AIPW estimators, marginal structural models)

09/22 - 10/13 – 6. Optimal Multiple Decision Treatment Regimes
- Characterization of an optimal regime (in terms of potential outcomes, observed data)
- Estimation of an optimal regime (Q-learning, direct search IPW/AIPW, backward iterative implementation, backward outcome weighted learning, A-learning, marginal structural models)
- Use of the R package DynTxRegime

10/13 - 10/22 – 7. Regimes Based on Time-to-Event Outcomes
- Basics of survival analysis
- Single decision statistical framework
- Single decision estimation for fixed and optimal regimes
- Multiple decision statistical framework

10/22 - 10/29 – 8. Sequential Multiple Assignment Randomized Trials (SMARTs)
- Design considerations
- Example SMARTs
- Power and sample size - simple comparisons
- Discussion

11/03 - 11/05 – 9. Statistical Inference for Treatment Regimes
- Nonsmoothness and statistical inference
- Inference for single decision regimes
- Inference for multiple decision problems (Q-learning)
- Power and sample size for SMARTs for optimal regimes

11/10 - 11/12 – Final Projects

**Class Evaluations for NCSU Students:** Online class evaluations will be available for students to complete before the end of the semester. Students will receive an email message directing them to a website where they can login using their Unity ID and complete evaluations. All evaluations are confidential; instructors will not know how any one student responded to any question, and students will not know the ratings for any instructors.

**Academic Integrity:** The instructor expects that students will abide by their institution’s policies on academic integrity. At NCSU, the University policy on academic integrity in the Code of Student Conduct Policy (POL 11.35.1) is available at [https://policies.ncsu.edu/policy/pol-11-35-01](https://policies.ncsu.edu/policy/pol-11-35-01). As
noted above, students may consult with one another on the homework, similar to how real researchers might consult with one another. However, students engaging in direct copying of the work or computer programs of fellow students will be considered in violation of policies on academic integrity.

**Students with Disabilities:** Reasonable accommodations will be made for students with verifiable disabilities. For NCSU students, to take advantage of available accommodations, students must register with Disability Services Office (DSO) at 2221 Student Health Services Building, Campus Box 7509, 515-7653; see [http://dso.dasa.ncsu.edu/](http://dso.dasa.ncsu.edu/) For more information on NCSU's policy on working with students with disabilities, please see the Academic Accommodations for Students with Disabilities Regulation (REG02.20.01) ([https://policies.ncsu.edu/regulation/reg-02-20-01/](https://policies.ncsu.edu/regulation/reg-02-20-01/)).