Syllabus for STAT 818R Fall 2026

Eric Slud 8/30/2026

This course introduces the mathematical statistics topic of Estimating Equations at a graduate level, using tools of advanced calculus and basic analysis. The topic arises in analyzing data via Generalized Linear Models, Biostatistics including Survival Analysis, Survey Inference, Missing Data methodology, Stochastic Regression, Longitudinal Models, Spatial Statistics, Robust Statistics or Quantile Regression models. The theory of Estimating Equations is primarily a frequentist large-sample subject area, connecting problems where an exact maximum likelihood analysis is infeasible or inconvenient either because one does not want to model all distributions in detail in order to define a likelihood or because the full likelihood is analytically intractable. That is, there can be modeling reasons or technical reasons to estimate parameters by solving an equation (usually defined by summing a function across observations) depending on both observations and parameters.

The basic idea is that, in data X_i, Y_i available for indviduals $i = 1, \ldots, n$ and independent identically distributed across individuals, an assumed relationship between a vector X_i of explanatory variables and a scalar or vector output Y_i may be expressed in the form $E(Q(X_i, Y_i, \theta)) = 0$ in terms of an unknown parameter for which the true value θ_0 is the unique solution, so that (subject to regularity conditions) the estimator $\theta^* = \theta$ solving $n^{-1} \sum_{i=1}^{n} Q(X_i, Y_i, \theta^*) = 0$ is consistent and asymptotically normal with mean θ_0 and variance approximately proportional to $1/n$.

The course will be run at Masters' level, like the Categorical Data class (STAT 770) that has often been taught in the past from the Agresti recommended text; that course material is more or less included in this course. The value of this class is in unifying the theoretical idea behind many practical applied models, including

Generalized Linear Models for Categorical Data (logistic regression, Poisson regression, etc.)

Longitudinal and Repeated-Measures Models

Quantile Regression Models

Weighted Models to compensate for Missing Data in Biostatistics and Surveys

The course will cover the models and basic theory and R computation and data examples.

Prerequisite: STAT 700, plus some computing familiarity. The level of the course will be similar to STAT 701.

Course requirements and Grading: there will be 6 graded homework sets (one every 2–2.5 weeks, split between theoretical and computational problems) which together will count 40% of the course grade, and a final project in the form of a 10–12 page paper presented orally on which there may be questions, that will serve as a final exam and will count 60% of the grade.

There will be no Required Course Text, but Recommended References and other course details including Homework assignments and due dates can be found on the course web-page https://www.math.umd.edu/∼slud/s818R.

Course Policies on Submitting Written Work

Homework and Final Project must be submitted electronically, on ELMS (Canvas), either in *.pdf format (scanned handwritten papers are acceptable if clearly legible) or *.doc format. They muist be submitted by the due date for full credit, unless for good reason I have granted you an extension.

Policies on use of Artificial Intelligence or LLM Internet Resources

As spelled out in University academic integrity policies: Generative AI tools may be used as a learning aid, for example to clarify concepts, review definitions, or help identify errors in your reasoning. However, they may not be used to generate solutions, proofs, derivations, or substantial portions of a homework or project paper submission. All submitted work must be written independently by you based on your own understanding. If you use a generative AI tool in a substantive way while working on an assignment, you must disclose how it was used. **Submitting AI-generated work as your own violates University academic integrity policies.**