

THEORETICAL STATISTICS Stat 210 B, Spring 2005, Tu Th 2-3:30,  
332 Evans

**Instructor:** Mark J. van der Laan, 108 Haviland Hall, 510-643-9866, laan@stat.berkeley.edu,  
www.stat.berkeley.edu/ laan.

**Teaching Assistant:** Dan Rubin, drubin@stat.berkeley.edu.

**Office hours:** Mark van der Laan: . Dan Rubin: .

**Book:** An important part of this course involves estimation methodology and applications covered in (Chapter 1, in particular) Mark J. van der Laan, James M. Robins (2002), Unified Methods for Censored Longitudinal Data and Causality, Springer Series in Statistics, Springer, New York. Therefore it is recommended to order this book.

**Hand Outs and Technical Reports:** We will also hand out some material in class. In particular, some relevant parts of Aad W. van der Vaart, Jon A. Wellner (1996), Weak Convergence and Empirical Processes, Springer Series in Statistics, Springer, New York. In addition, we will refer to some of our publications/technical reports available on the web [www.bepress.com/ucbbiostat](http://www.bepress.com/ucbbiostat). Possibly, some lecture notes will be posted on my website [www.stat.berkeley.edu/ laan](http://www.stat.berkeley.edu/laan).

**Evaluation:** The evaluation of each student is based on attendance, a few homework assignments, midterm, and final exam.

**Topic I Basic Empirical Process Theory:** Reference: Empirical Processes, van der Vaart, Wellner (1996), Chapter 2.

1. Identically and independently distributed observations of a random variable (the experimental unit), and its corresponding empirical distribution.
2. Empirical process indexed by a class of functions (of the experimental unit).
3. Weak convergence of an empirical process in function space endowed with sigma-algebra. Equivalence of weak convergence of empirical process with convergence of finite dimensional distributions and uniform continuity of the sample paths w.r.t. variance norm. Donsker class of functions. Consistency of empirical process in function space. Glivenko-Cantelli class of functions.

4. Characterization of Donsker classes in terms of entropy and bracketing numbers. Examples of Donsker classes. Permanence of Donsker classes.
5. Some empirical process inequalities: Bernstein's inequality.

**Topic II Establishing consistency and asymptotic linearity/weak convergence of estimators**

Possible reference: Empirical Processes, van der Vaart, Wellner (1996), Chapter 3.

1. The extended continuous mapping theorem.
2. Asymptotic linearity of an estimator. The influence curve of an asymptotically linear estimator.
3. The functional derivative of an estimator viewed as a function of the empirical distribution, and its relation to the influence curve. Functional delta method.
4. M-estimators, that is, estimators which are defined as solution of an estimating equation. The analysis of M-estimators which are smooth functions of the empirical distribution, and establishing its influence curve.

**Topic III Estimating functions for smooth parameters:** Reference: Chapter 1 (and more, if you wish) of van der Laan, Robins (2002), Unified Methods for Censored Longitudinal Data and Causality, Springer.

1. Statistical framework for independent and identically distributed observations  $O_1, \dots, O_n \sim P_0$ : Data, Model, Parameter of Interest.
2. Efficiency Theory for regular asymptotically linear estimators: influence curve of an estimator, class of parametric one-dimensional submodels, scores in Hilbert space  $L_0^2(P_0)$ , nuisance scores, tangent space, nuisance tangent space, pathwise derivative of the parameter of interest w.r.t to this class of 1-d submodels, gradients, canonical gradient/efficient influence curve.
3. Orthogonal complement of nuisance tangent space and corresponding class of estimating functions for parameter of interest.
4. Relating Full Data estimating functions and Censored Data estimating functions.

5. Examples of classes of estimating functions in regression, censored data regression, multiplicative intensity models, and more.
6. Robustness of unbiasedness of Estimating Functions w.r.t. misspecification of nuisance parameter: First order robustness for variation independent nuisance parameter of estimating function, complete robustness for variation independent convex linear nuisance parameter, double robustness.
7. Locally efficient estimation based on estimating functions.

**Topic IV Unified Loss Based Estimation/Learning:** References: Technical reports on [www.bepress.com/ucbbiostat](http://www.bepress.com/ucbbiostat).

Mark J. van der Laan, Sandrine Dudoit, and Aad W. van der Vaart, "The Cross-Validated Adaptive Epsilon-Net Estimator" (February 26, 2004). U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 142.

<http://www.bepress.com/ucbbiostat/paper142>

Mark J. van der Laan and Sandrine Dudoit, "Unified Cross-Validation Methodology For Selection Among Estimators and a General Cross-Validated Adaptive Epsilon-Net Estimator: Finite Sample Oracle Inequalities and Examples" (November 26, 2003). U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 130.

<http://www.bepress.com/ucbbiostat/paper130>

Sandra E. Sinisi and Mark J. van der Laan, "Loss-Based Cross-Validated Deletion/Substitution/Addition Algorithms in Estimation" (March 3, 2004). U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 143.

<http://www.bepress.com/ucbbiostat/paper143>

Annette Molinaro and Mark J. van der Laan (November 9, 2004) Deletion/Substitution/Addition Algorithm for Partitioning the Covariate Space in Prediction

Yue Wang and Mark J. van der Laan, "Data Adaptive Estimation of the Treatment Specific Mean" (October 20, 2004). U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 159.

<http://www.bepress.com/ucbbiostat/paper159>

1. Statistical Framework in terms of parameter-specific loss function and or risk function

2. Road map for constructing a general data adaptive estimator: parameter space, sieve on parameter space, sieve on nuisance parameter of loss function, subspace specific minimum empirical risk estimators, cross-validation to select among candidate estimators.
3. Formal finite sample and asymptotic results (minimax adaptivity) for cross-validation selector and for general adaptive estimator based on epsilon-net sieves
4. Deletion/substitution/addition (DSA) algorithms for computing the subspace specific minimum empirical risk estimator.
5. Examples.

**Topic Va Unified Estimating Function Based Cross-Validation:** Reference: van der Laan, Rubin (2005), Technical report will appear on [www.bepress.com/ucbbiostat](http://www.bepress.com/ucbbiostat).

**Topic Vb Unified Estimating Function Based Learning:** Estimating function based learning, analogue to road map in Topic IV.

**Future Book:** Unified Minimax Adaptive Learning, in progress.