

Sündüz Keleş

2109 Shattuck Avenue, #305
Berkeley, CA 94704.
(510) 704-9142

sunduz@stat.berkeley.edu
www.stat.berkeley.edu/~sunduz

EDUCATION

UC Berkeley, Berkeley, CA.

Ph.D. in Biostatistics, May 2003 (expected).

Dissertation Title: *Statistical methods for detecting cis-regulatory motifs in biological sequences and Estimation of the bivariate survival function with right censored data structures.*

Committee: Mark J. van der Laan (chair), Peter Bickel, Michael Jordan.

UC Berkeley, Berkeley, CA.

M.A. in Biostatistics, May 2000.

Bilkent University, Ankara, Turkey.

B.Sc. in Industrial Engineering, May 1998.

PUBLICATIONS

S. Keleş, M.R. Segal (2002), Residual-based tree-structured survival analysis. *Statistics in Medicine*, 21: 313-326.

S. Keleş, M.J. van der Laan, M. B. Eisen (2002), Identification of regulatory elements using a feature selection method. *Bioinformatics*, 18: 1167-1175.

S. Keleş, M.J. van der Laan, J.M. Robins (2002), Estimation of the bivariate survival function with right censored data structures. *Invited paper for the Volume on Survival Analysis of the book series Handbook of Statistics*, August 2002.

M. Pavlic, **S.Keleş**, M.J. van der Laan, S. Butler (2002), Recurrent Events analysis in the presence of time dependent covariates and dependent censoring. *Revised for the Journal of Royal Statistical Society Series B*, November 2002.

S. Keleş, M.J. van der Laan, S. Dudoit, B. Xing, M.B. Eisen (2002), Detecting regulatory motifs with entropy constraints in DNA sequences. *Submitted to Genome Biology*, November 2002.

S. Keleş, M.J. van der Laan, S. Dudoit (2002), Asymptotically optimal model selection method for regression on censored outcomes. *Submitted to Bernoulli*, December 2002.

S. Keleş, M.J. van der Laan, C. Vulpe (2002), Finding combinatorial regulatory motifs using gene expression data. *In preparation.*

RESEARCH INTERESTS

Computational biology, bioinformatics, analysis of complex longitudinal data, survival analysis, multivariate statistical methods.

RESEARCH
EXPERIENCE

Research Assistant, Division of Biostatistics, UC Berkeley, May 2000-present.
Supervisor: Mark J. van der Laan.

- Developed two statistical methods for finding transcription factor binding sites. In the first method, modeled the gene expression from microarray data as a function of the non-coding sequences and formulated the biological question of interest as a feature selection problem where relevant features are selected with a stepwise cross-validation method. As a second method, developed a mixture model of non-coding sequence data that employs specific constraints on the entropy structure of the binding site. Emphasis on identifying binding sites with structural constraints on DNA-transcription factor. Collaborated with Professor **M.B. Eisen** and his lab at UC Berkeley in both of the projects.

- Worked on the estimation of the bivariate survival function with generalized right censored data structures (including time dependent/independent covariate processes and subject to informative censoring). Proposed a generalized Dabrowska's estimator that remains consistent under informative censoring using an estimating function approach. Applied a double robust estimation methodology to this generalized estimator to improve efficiency and robustness.

Research Assistant, Department of Nutritional Sciences and Toxicology, UC Berkeley, May 2001-May 2002. Supervisors: Mark J. van der Laan, Chris Vulpe.

- Analyzed non-coding sequence data of the yeast *Saccharomyces cerevisiae* genes that are involved in iron-copper metabolism. Developed a tool that integrates enumerative motif finding methods and logic regression to find discriminative and combinatorial non-coding sequence elements of up and down regulated genes. Collaborated with **Vulpe Lab** at UC Berkeley.

Research Assistant, Division of Epidemiology and Biostatistics, UCSF, May 1999-May 2000. Supervisor: Mark R. Segal.

- Developed an analytical relationship between the log-rank split criteria and martingale residual sum-of-squares split criteria that are used in tree structured survival analysis.

PRESENTATIONS

Identification of regulatory motifs using gene expression data.

Invited Talk, Applications of Statistics to Molecular Biology and Genetics Seminar, October 26, 2001. UC Berkeley, CA.

Invited Talk, Genomics and Computational Biology Seminar Series, November 7, 2001. UC Berkeley, CA.

Identification of regulatory elements using a feature selection method.

Contributed talk, WNAR/IMS Meeting, June 23-26, 2002. UCLA, Los Angeles, CA.

Estimation of the bivariate survival function under informative censoring with time-dependent covariates.

Contributed talk, Joint Statistical Meeting, August 11-15, 2002. New York, NY.

Detecting structured motifs in DNA sequences.

Invited talk, Applications of Statistics to Molecular Biology and Genetics Seminar, September 26, 2002. UC Berkeley, CA.

Sündüz Keleş

COMPUTING SKILLS	Proficient in C, R/S-plus, Perl, Matlab, Fortran, Pascal. Experienced with SAS, Html. Proficient in working on Unix/Linux, Windows platforms.
AWARDS AND HONORS	Biostatistics Block Grant Award, Fall 1999-Spring 2000. Biostatistics Block Grant Award, Fall 1998-Spring 1999. Full Scholarship by Bilkent University, 1994-1998. Ranked in top 200 among 1 million participants in the nationwide university entrance exam, Turkey, 1994.
TEACHING EXPERIENCE	Teaching Assistant , Division of Biostatistics, UC Berkeley, Berkeley, CA. Lectured for Statistical Techniques in Computational Biology course, Fall 2001. Volunteer Tutor , Ankara, Turkey. Gave lectures to 30+ high school students on word processor and spread sheets, Fall 1997.
SERVICES	Refereed for <i>Biometrics</i> , <i>Biostatistics</i> , and <i>Bioinformatics</i> 2002. Member of the website committee of UC Berkeley Designated Emphasis in Genomics and Computational Biology.

REFERENCES

- 1. Mark J. van der Laan**, Professor of Biostatistics and Statistics. (Thesis advisor)
School of Public Health, Division of Biostatistics, UC Berkeley.
Earl Warren Hall #7360,
Berkeley, CA 94720-7360.
(510) 643-9866
(510) 643-5163 (fax)
laan@stat.berkeley.edu
- 2. Mark R. Segal**, Professor in Residence, Division of Biostatistics; Vice-Chair,
Department of Epidemiology & Biostatistics, UCSF.
500 Parnassus Avenue, 420 MU-W,
San Francisco, CA 94143-0560.
(415) 476-4553
(415) 476-6014 (fax)
mark@biostat.ucsf.edu
- 3. Michael B. Eisen**, Scientist, Life Sciences Division, Lawrence Berkeley National
Lab, and Adjunct Assistant Professor, Department of Molecular and Cellular Biology,
UC Berkeley.
Mailstop 84-171 One Cyclotron Road,
Berkeley, CA 94720.
(510) 486-5214
(786) 549-0137 (fax)
mbeisen@lbl.gov
- 4. Peter Bickel**, Professor of Statistics.
Department of Statistics, UC Berkeley.
367 Evans Hall,
Berkeley, CA 94710-3860.
(510) 642-2781
(510) 642-7892 (fax)
bickel@stat.berkeley.edu
- 5. Nicholas P. Jewell**, Professor of Biostatistics and Statistics, Chair.
School of Public Health, Division of Biostatistics, UC Berkeley.
Earl Warren Hall,
Berkeley, CA 94720-7360.
(510) 642-3241
(510) 643-5163 (fax)
jewell@stat.berkeley.edu