OMB No. 0925-0001/0002 (Rev. 08/12 Approved Through 8/31/2015)

BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. DO NOT EXCEED FIVE PAGES.

NAME: Bradley Efron

eRA COMMONS USER NAME (credential, e.g., agency login): EFRON, BRADLEY

POSITION TITLE: School of Humanities and Sciences Max H. Stein Professor of Statistics and Biostatistics

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)

| INSTITUTION AND LOCATION | DEGREE(if applicable) | Completion DateMM/YYYY | FIELD OF STUDY |
| --- | --- | --- | --- |
| California Institute of Technology | BS | 1960 | Mathematics |
| Stanford University, California | MS | 1962 | Statistics |
| Stanford University, California | Ph.D. | 1964 | Statistics |

# A. Personal Statement

Efron’s 2011 book on large-scale inference is based on his published papers concerning the statistical analysis of massive biomedical data sets, with an emphasis on empirical Bayes methods. Since then his research has focused on the relationship between Bayes, empirical Bayes, and frequentist applications, particularly the role of the bootstrap as a connecting methodology. His paper, “Estimation and accuracy after model selection,” was selected by the *Journal of the American Statistical Association* Theory and Methods section for presentation and discussion at the 2014 Joint Statistical Meetings in Boston. His main work on Cox regression is “The efficiency of Cox's likelihood function for censored data” (1977, *JASA* **72**, 557–565). It includes the so-called “Efron approximation” for handling cases with tied survival times, so-named in *Modeling Survival Data* by Therneau & Grabbsch (2000).

# B. Positions and Honors

**Positions**

1966--1971 Assistant and Associate Professor of Statistics, Stanford University

1968--1969 Associate Editor, *Journal of the American Statistical Association*

1969--1972 Theory and Methods Editor, *Journal of the American Statistical Association*

1972-- Professor of Statistics, and Health Research and Policy, Stanford University

1976--1979 Chairman, Department of Statistics, Stanford University

1967--1968 Visiting Lecturer, Dept. of Statistics, Harvard University

1971--1972 Visiting Scholar, Dept. of Mathematics, Imperial College, England

1979--1980 Visiting Professor, Dept. of Statistics, UC Berkeley

1981-- Chairman, Mathematical and Computational Sciences Program, Stanford University

1985--1986 Chairman, Overseers Committee, Department of Statistics, Harvard University

1987--1988 President, Institute of Mathematical Statistics

1987-1990 Associate Dean, School of Humanities and Sciences, Stanford University

1991--1994 &

1998--1999 Chairman, Department of Statistics, Stanford University

1994 & 1997 Chairman, University Advisory Board, Stanford University

1998--1999 Chairman, Stanford University Faculty Senate

2004 President, American Statistical Association

2006--2012 Founding Editor, *Annals of Applied Statistics*

**Honors**

Rietz Lecturer, Annual Meeting of IMS, August 1977

Ford Prize, Mathematical Association of America, 1978

Wald Lecturer, Annual Meeting of the IMS, August 1981

``Outstanding Statistician of The Year'', ASA, Chicago Chapter, 1981

MacArthur Prize Fellow, January 1983

American Academy of Arts and Sciences, 1983

National Academy of Sciences, 1985

Endowed Chair, Max H. Stein Professor of Statistics and Biostatistics, December 1987

Wilks Medal, American Statistical Association, August 1990

Honorary Doctor of Science Degree, University of Chicago, June 1995

Fisher Prize, COPSS, July 1996

Doctor Honoris Causa, Universidad Carlos III de Madrid, Spain, 1998

Parzen Prize for Statistical Innovation, 1998

Doctor Honoris Causa, University of Oslo, Norway, 2002

C.R. and Bhargarvi Rao Prize, The Pennsylvania State University, 2003

Noether Prize, American Statistical Association, 2006

National Medal of Science, 2005 (awarded 2007)

Distinguished Alumni Award, California Institute of Technology, 2010

Guy Medal in Gold, Royal Statistical Society, 2014

# C. Contribution to Science:

1. BOOTSTRAP METHODS. The bootstrap is a general method for assessing the accuracy of a statistic computed from an observed set of data. It substitutes computation for the theoretical approximations familiar in applied work. The payoff is that the bootstrap estimates of accuracy can be carried out for almost any problem, no matter how complicated the underlying probability model may be. Since my original 1979 paper I have pursued the underlying theory supporting the calculations, with extensions to confidence intervals and measures of prediction error.
	1. Bootstrap methods: Another look at the jackknife (1979). *Ann. Statist*. **7**, 1–26.
	2. Better bootstrap confidence intervals (1987). *JASA* **82**, 171–200 with discussion and Rejoinder.
	3. Bootstrap confidence intervals (with T. DiCiccio) (1996). *Statist. Sci*. **11**, 189–212.
	4. The bootstrap and Markov chain Monte Carlo (2011). *J. Biopharm. Statist*. **21**, 1052–1062. doi:10.1080/10543406.2011.607736
2. LARGE-SCALE TESTING. Modern scientific technology produces massive data sets, often with thousands of hypothesis tests to perform at the same time. After introducing “local false discovery rates” for this task, I have pursued both their theory and application in applied practice.
	1. Empirical Bayes analysis of a microarray experiment (with R. Tibshirani, J.D. Storey and V. Tusher) (2001). *JASA* **96**, 1151–1160.
	2. Large-scale simultaneous hypothesis testing: The choice of a null hypothesis (2004). *JASA* **99**, 96–104.
	3. Are a set of microarrays independent of each other? (2009) *Ann. Appl. Statist*. **3**, 922–942. doi:10.1214/09-AOAS236
	4. Correlated *z*-values and the accuracy of large-scale statistical estimates (2010). *JASA* **105**, 1042–1069 with discussion and Rejoinder. doi:10.1198/jasa.2010.tm09129
3. EMPIRICAL BAYES ESTIMATION AND TESTING. Data sets consisting of a large number of small parallel problems, for instance the simultaneous testing of significance of genes in a microarray study, are amenable to empirical Bayes analysis; that is, one can approximate Bayesian results for any one problem using the data from all the “other” problems. This theory emerged in the 1950s in the work of C. Stein and also of H. Robbins. I have tried to put the theory into practical form for modern application, as well as understanding its basis in classical Bayesian and frequentist philosophy.
	1. Stein's estimation rule and its competitors – An empirical Bayes approach (with C. Morris) (1973). *JASA* **68**, 117–130.
	2. Estimating the number of unseen species: How many words did Shakespeare know? (with R. Thisted) (1976). *Biometrika* **63**, 435–447.
	3. *Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing, and Prediction* (2010). Institute of Mathematical Statistics Monographs **I**, Cambridge University Press, Cambridge.
	4. Two modeling strategies for empirical Bayes estimation (2014). *Statist. Sci*. **29**, 285–301. doi:10.1214/13-STS455
	5. Empirical Bayes deconvolution estimates (2016). *Biometrika* **103**, 1–20. doi:10.1093/biomet/asv068
4. ERROR RATES FOR PREDICTION ALGORITHMS. Given a training set of covariate vectors and their associated responses, a prediction rule is constructed; having observed a new covariate vector, the rule predicts what the response might be. How accurate are the rule's predictions? Two main theories have been put forward, cross-validation and “Cp” estimates. I have pursued the theory and application of both methods, and their relationship, in a series of papers.
	1. Estimating the error rate of a prediction rule: Improvement on cross-validation (1983). *JASA* 78, 316–331.
	2. How biased is the apparent error rate of a prediction rule? (1986). *JASA* **81**, 461–470.
	3. Improvements on cross-validation: The :632+ bootstrap method (with R. Tibshirani) (1997). *JASA* **92**, 548–560.
	4. Least angle regression (with T. Hastie, I. Johnstone and R. Tibshirani) (2004). *Ann. Statist*. **32**, 407–499 with discussion and Rejoinder. doi:10.1214/009053604000000067
	5. The estimation of prediction error: Covariance penalties and cross-validation (2004). *JASA* **99**, 619–642 with discussion.
5. BAYESIAN AND FREQUENTIST INFERENCE. There are two main statistical philosophies, Bayes and frequentist, often in tension with each other in applied practice. Understanding their strengths, weaknesses, and relationship has taken on new importance in our era of large-scale inference. I have written a series of papers investigating how the philosophies play out in the real world.
	1. Why isn't everyone a Bayesian? (1986). *Amer. Statist*. **40**, 1–11.
	2. Statistical data analysis in the computer age (with R. Tibshirani) (1991). *SCIENCE* **253**, 390–395.
	3. R.A. Fisher in the 21st century (1998). *Statist. Sci*. **13**, 95–122 with discussion and Rejoinder.
	4. Bayesians, frequentists, and physicists (2003). *Proc. PHYSTAT2003*, 17–28.
	5. A 250-year argument: Belief, behavior, and the bootstrap (2013). *Bull. Amer. Math. Soc*. **50**, 129–146. doi:10.1090/S0273-0979-2012-01374-5

# D. Research Support

## *Ongoing Research Support*

NSF DMS 1208787 (Efron) 07/01/2016–06/30/2019

Statistical Theory and Methodology

The aim of this research is to develop theoretical ideas in mathematical statistics and probability. The more theoretical work on scatterplot smoothing was partially supported by this grant. The NIH grant and this NSF grant jointly supported Efron's work on model selection.

NIH UL1 TR00108503 (Greenberg, Harry B.) 09/26/2013–04/30/2018

Spectrum Stanford Center for clinical and Translational Research and Education

This is a trans-institutional CTSA to transform clinical and translational research and education at Stanford. It consists of 12 separate programs, each focused on ensuring that basic discoveries get translated into improvements in human health and wellbeing.

NIH 1 UH2 TR000902 (Matin, A.C.) 08/01/2013–07/31/2018

HER2-targeted exosomal delivery of therapeutic mRNA for enzyme pro-drug therapy

To specifically deliver a prodrug regimen to HER2+ve breast cancer for treating it with no or minimal side effects.

## *Completed Research Support*

NSF DMS 1208787 (Efron) 07/01/2012–06/30/2016

Statistical Theory and Methodology

The aim of this research is to develop theoretical ideas in mathematical statistics and probability. The more theoretical work on scatterplot smoothing was partially supported by this grant. The NIH grant and this NSF grant jointly supported Efron's work on model selection. Role: PI

NIH 1R01 MH10178203 (Sabatti, Chiara) 08/01/2013–07/31/2016

Genetic Regulation of Gene Expression and its Impact on Phenotypes

This project will develop statistical methodologies to identify genetic determinants of differential gene expression across tissues; we will rely on meta-analysis and on new multiple comparisons correction procedures. We will apply the methods to the analysis of GTEx data as well as datasets collected in the study of bipolar disorder and dyslipidemia.

NIH 8R37 EB002784 (Efron) 02/01/2010–01/31/2016

Adaptation of New Statistical Ideas for Medicine

The primary goal for this grant is the practical analysis of statistical power for microarray studies. Power here refers to the ability of the study to identify individual genes that are involved in differential Expression. Role: PI

NSF DMS 1208787 (Efron) 07/01/2012–06/30/2015

Statistical Theory and Methodology

The aim of this research is to develop theoretical ideas in mathematical statistics and probability. The more theoretical work on scatterplot smoothing was partially supported by this grant. The NIH grant and this NSF grant jointly supported Efron's work on model selection. Role: PI