The Rising of Academic Statistics and Biostatistics

(with some focus on New England)

Alan Agresti

Professor Emeritus, University of Florida

part-time New Englander

Outline

- My talk is based on a book I recently edited with Xiao-Li Meng:

  Strength in Numbers: The Rising of Academic Statistics Departments in the U.S. (published by Springer)

- Evolution of Statistics and Biostatistics departments in the U.S.

- Two department vignettes:
  A failure (Princeton) and a success (NC State)

- Departments in New England

Why this book project?

Recent department birthdays (*AmStat News*, *IMS Bulletin*):

Iowa State Statistics: 75

North Carolina State Statistics: 60

Virginia Tech Statistics: 60

U of North Carolina Biostatistics: 60

U of Michigan Biostatistics: 60

Harvard Statistics: 50

Wisconsin Statistics: 50

Agresti and Meng (September 2009):
Is memory of early days of many departments disappearing?
Let's try to capture some history, before it may be too late.

Scope? Invited chapters from each of 40 U.S. Statistics and Biostatistics departments in existence by mid 1960s

Statisticians and Statistics courses in academia

- Departments not needed for Statistics instruction: Available by 1890 at universities such as Columbia, Johns Hopkins, Michigan, MIT, Pennsylvania, Virginia, Wisconsin, and Yale.
- e.g., U. of Wisconsin 1890: George Comstock publishes text based upon his course “Method of Least Squares,” targeted to students of physics, astronomy, engineering
- Through first half of twentieth century, statisticians mainly worked in isolation, in departments of Economics, Business, Mathematics, Education, Psychology

Advanced statistics courses?

1920s: Center of advanced statistics instruction moves to departments of mathematics, following books such as

G. Udny Yule: *Intro. to the Theory of Statistics* (1911)

Carl West: *Intro. to Mathematical Statistics* (1918)

Henry Rietz: *Handbook of Mathematical Statistics* (1924) and *Mathematical Statistics* (1927).

(Yule and West available at www.forgottenbooks.org)

Who should teach statistics courses?

Harold Hotelling (1940, 1949): Statistics well deserving of its own department. Academia not well served by having Statistics taught in various departments by teachers inadequately prepared in statistical theory.

“The task of leading the blind must not be turned over to the blind.”

Hotelling also argued against Statistics being organized under Department of Mathematics.

Who should teach statistics courses?

Hotelling: “The teacher of statistics must have a genuine sympathy and understanding for applications, and these are not well possessed by a great many pure mathematicians. ... specialists will be suspicious that courses in statistics given by a department consisting largely of pure mathematicians are unsuitable for their purposes. The result is likely to be a sabotaging of attempts at centralization, the different departments reverting to the old and ultimately objectionable system of teaching their own separate courses in statistical methods.”

George Box (2013): “A serious mistake has been made in classifying statistics as part of the mathematical sciences. Rather it should be regarded as a catalyst to scientific method instead.”

Pre World War II: Evolution of Statistics/Biostatistics

- 1918: Johns Hopkins Dept of Biometry and Vital Statistics
- 1922: Harvard (HSPH) Department of Vital Statistics
- 1930: Buffalo Department of Statistics and Insurance
- 1931: U. of Pennsylvania Department of Economic and Social Statistics
- 1933: Iowa State Statistical Laboratory
- 1935: George Washington Univ. Statistics Department
- 1938: Statistical Laboratory, U. of California at Berkeley
- 1941: North Carolina State Dept. of Experimental Statistics

Outside U.S.: University College, London, 1911 (Karl Pearson)

Post World War II development

Statistics departmental creation explodes by mid-1960s:

Columbia 1946, North Carolina 1946, Cornell Biometrics 1947, Purdue 1947, Stanford 1948, Virginia Tech 1948, Chicago 1949, Michigan Public Health Statistics 1949, Pittsburgh Biostatistics 1949, UNC Biostatistics 1949, Rutgers 1952, Michigan State 1955, Harvard 1957, Oregon State 1957, Minnesota 1958, Florida State 1959, Kansas State 1959, Wisconsin 1960, SMU 1961, Connecticut 1962, Florida 1962, Iowa 1962, Texas A & M 1962, Missouri 1963, Yale 1963, Georgia 1964, Washington Biostatistics 1964, Princeton 1965

By 1970 *AmStat News* listing of colleges and universities offering degrees in Statistics, 99 departments or programs existed with “Statistics” or “Biostatistics” in the title.

Bumps in the road

Departmental creation and building did not always come smoothly. Examples:

- In 1940 at Wisconsin, visiting lecturer Milton Friedman prepares report proposing Department of Statistics and is invited by administration to implement it, but department not formed until 1960, when George Box invited to come from Princeton to do so.
- Strong disagreements about vision sometimes resulted in splits or key people disengaging, or even disbanding of the department (Princeton and Buffalo in 1980s).

Departmental vignettes: Princeton Statistics

(by Dave Hoaglin and Karen Kafadar)

- Sam Wilks arrives 1933 in Math Dept., eventually creates Section of Mathematical Statistics
- After collaborating with Wilks at Fire Control Research Office during WWII, John Tukey returns to Princeton, dividing time with AT&T.
- Founded Statistical Techniques Research Group with George Box as director (1956-1959) and many visitors.
- PhD recipients include T. W. Anderson, P. Billingley, D. Brillinger, A. Dempster, D. Fraser, D. Freedman, L. Goodman, J. Hartigan, P. Meier, F. Mosteller, D. Wallace.

Princeton: A short-lived department

- Wilks dies 1964, department established 1965 with Tukey as chair (part-time until 1970), succeeded by Geoff Watson and Peter Bloomfield
- Department had impressive list of visitors, highlights include 1970-1971 robustness year, Tukey's 1977 *Exploratory Data Analysis*.
- With retirements (Tukey 1985), moves (Bloomfield to NC State 1982 and Watson to Mathematics 1985), administration effectively disbands department in 1985.

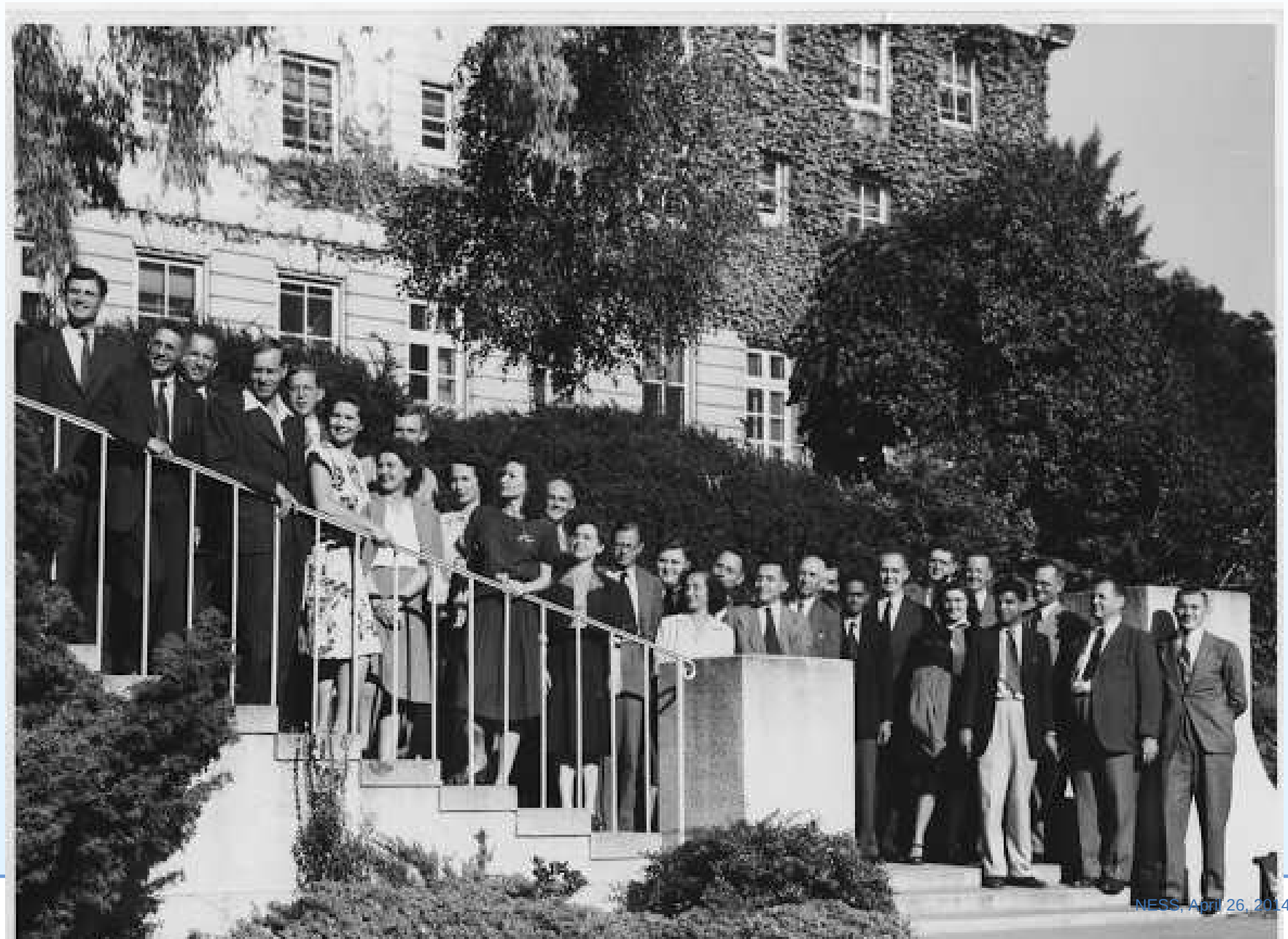
Departmental vignettes: NC State Statistics

(by Tom Gerig)

- 1940: Search committee to appoint director to develop center of Statistics includes George Snedecor, who shows list of suitable candidates to colleagues. Gertrude Cox responds “I don’t see my name on there.” Snedecor attaches postscript, “If you are willing to consider a woman, I know of none better qualified than Gertrude M. Cox.”
- In January 1941, Cox is appointed professor and head of Department of Experimental Statistics (and the first female professor of any rank at NC State).
- Immediately establishes consulting service and summer courses with faculty such as Harold Hotelling, R. A. Fisher, William Cochran.

NC State Statistics summer session, 1946

At Berkeley Statlab, summer session 1947



Virginia Polytechnic Institute summer session 1947

Departmental vignettes: Gertrude Cox at NC State

- Hires in 1940s included William Cochran, who serves as associate director and builds graduate program in Statistics.
- 1944: Cox forms and directs state-funded “Institute of Statistics,” which expands to include a theoretical statistics department at UNC, with Hotelling as founding head.
- Cox serves as department chair until 1949, Institute director until 1960, when helps establish Research Triangle Institute, heads its Statistics Research Division until her retirement in 1965.
- Founding editor of *Biometrics* in 1945, publishes *Experimental Design* with Cochran in 1950, serves as President of ASA in 1956.

William Cochran, Gertrude Cox, Harold Hotelling, 1944

NC State Statistics – Later years

- Under heads Jackson Rigney (1949-1963), David Mason (1963-1981), Dan Solomon (1981-1993), Tom Gerig (1994-2002), Sastry Pantula (2002-2010), Montserrat Fuentes (2011-), faculty grows, with strengths in design and linear models, econometrics, time series, nonparametrics, biomathematics, biostatistics (e.g., Dennis Boos, Peter Bloomfield, David Dickey, Andrew Gallant, Butch Tsiatis, Marie Davidian).
- 1966-1976: SAS project initiated by Tony Barr with software development by James Goodnight, moving off campus in 1976 and establishing as SAS Institute, Inc.
- Graduate students: 57 in 1980, 140 in 2012
- *Biometrics* editorship and ASA Presidency tradition continues recently with Marie Davidian.

NC State faculty, 2004

New England Statistics and Biostatistics Departments

Many of you know more than I do, so please add your comments and anecdotes later.

Departmental vignettes: Univ. of Connecticut Statistics

(by D. Dey, N. Mukhopadhyay, L. Kuo, and M.-H. Chen)

- Robert Riffenburgh hired from Virginia Tech in 1962 to develop Statistics program. Department consulting service initiated in 1963.
- Hirings, by chair (up to current, Joseph Glaz) include:
  - Riffenburgh (1962-1968): Alan Gelfand, Uwe Koehn, Gottfried Noether, Harry Posten
  - Gottfried Noether (1968-1982): Tim Killeen, Joseph Glaz
  - Uwe Koehn (1982-1987, 1991-1996): Dipak Dey, Nitis Mukhopadhyay, Lynn Kuo, Richard Vitale
  - Nitis Mukhopadhyay (1987-1990): Nalini Ravishanker, Suman Mazumdar,
  - Dipak Dey (1997-2011): Vladimir Pozdnyakov, Ming-Hui Chen, Zhiyi Chi, Ofer Harel, Jun Yan

Connecticut Statistics in the 60s

STATISTICS DEPARTMENT STAFF

Robert H. Riffenburgh, Professor and Head
B.S., M.S. (William and Mary), Ph.D. (Va. Polytechnic Inst.)
Areas of specialization: mathematical models, multivariate analysis, Markov processes.
Activities: administration, research, teaching, consulting.
Experience: academic, industrial, governmental.

Joseph J. Lucas, Associate Professor; Experiment Station Biometrician
B.S., M.S. (Univ. of Calif.), Ph.D. (Wash. State Univ.)
Areas of specialization: genetic statistics, statistics applied to agriculture.
Activities: consulting, teaching, research.
Experience: academic, professional consulting.

Harry O. Posten, Associate Professor
B.S. (Central Conn. State College, M.S. (Kansas State Univ.), Ph.D. (Va. Polytechnic Inst.)
Areas of specialization: multivariate analysis, statistical robustness.
Activities: teaching, research, consulting.
Experience: academic, industrial.

Earl J. Bell, Assistant Professor
B.S., M.S., Ph.D. (Univ. of California)
Areas of specialization: operations research, mathematical programming.
Activities: teaching, research.
Experience: academic.

Helen Jean Thiebaux, Assistant Professor
B.A. (Reed College), M.A. (Univ. of Oregon), Ph.D. (Stanford Univ.)
Areas of specialization: probabilistic evaluation of assumptions, decision theory.
Activities: teaching, research.
Experience: academic.

Derrick S. Tracy, Assistant Professor
B.S., M.Sc. (Lucknow Univ.), M.S., Ph.D. (Univ. of Michigan)
Areas of specialization: sampling and distribution theory.
Activities: teaching, research, consulting.
Experience, academic, governmental.

Samuel Zahl, Assistant Professor
B.S., M.S. (Univ. of Chicago), Ph.D. (Harvard Univ.)
Areas of specialization: distribution theory, decision theory, operations research.
Activities: teaching, research, consulting.
Experience: academic, industrial, governmental.

Lecturers in Statistics

Connecticut's varied contributions

- NESS began at U. Connecticut in 1987
- Pfizer Colloquium Series started 1979 with videotapes of distinguished statisticians, including Neyman, Cramer, Rao, Chernoff, Deming, Box, Kempthorne, ... (ASA archives)
- Meetings hosted or originated by department include International Workshop in Applied Probability, International Chinese Statistical Association Applied Statistics Symposium, International Indian Statistical Association, International Workshop in Sequential Methodologies
- PhD graduates include David Salsburg, Chris Tsokos, Brad Carlin, Bani Mallick, Sudipto Banerjee
- Distinguished research particularly in Bayesian statistics (e.g., Gelfand, Chen, Dey, Ravishanker, Kuo) and applied probability (e.g., Glaz, Chi, Mukhopadhyay, Pozddnyakov)

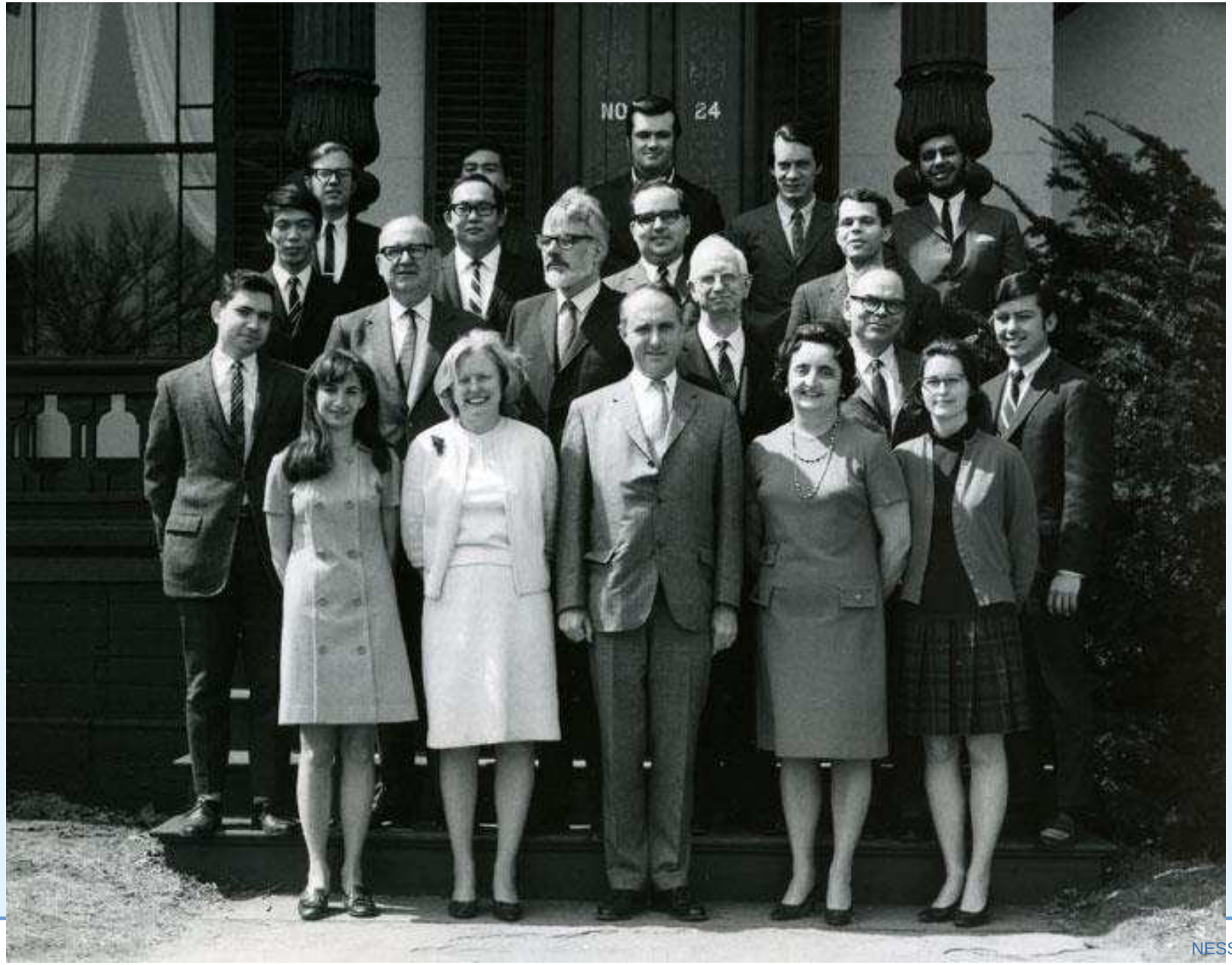
Connecticut Statistics photo (2011)

Departmental vignettes: Yale University Statistics

(by John Hartigan)

- PhD awarded at Yale in 1876 to Mansfield Merriman for thesis on method of least squares (Stigler 1978).
- 1896 course (Economics 21, still taught) by Irving Fisher on use and abuse of statistics. Fisher, co-founder of Econometric Society, taught until retirement in 1935.
- University Committee on Statistics, Statistics Laboratory, established in 1953.
- Statistics department founded in 1963, Francis Anscombe coming from Princeton as chair.
- Early faculty included Chester Bliss (biometry, probit model), Jimmie Savage (Bayes), John Hartigan (classification).

Yale Statistics photo (1968)



Yale: Small department survives and then thrives

- Jimmie Savage's sudden death (1971) could have ended department, which relied heavily on visitors until hirings in mid-1970s including Richard Savage, David Pollard.
- Following retirements and losses, in 1991 only two tenured (Hartigan, Pollard) and two non-tenured faculty. Andrew Barron's hiring provides new energy, directions.
- Distinguished alumni include Don Berry, Deborah Nolan, Richard Olshen, William DuMouchel, Edward Carlstein, Gary Oehlert
- Undergraduate major approved in 2009.
- In 2014, four tenured faculty (Barron, Pollard, Joseph Chang, Harrison Zhou), six non-tenured faculty (Lisha Chen, John Emerson, Mokshay Madiman, Sahand Negahban, Jonathan Reuning-Scherer, Jing Zhang).
- Division of Biostatistics in Public Health shares faculty.

Yale Statistics photo (2003)

Departmental vignettes: Harvard Statistics

(by Xiao-Li Meng)

- Department established in 1957 under Frederick Mosteller, from Dept of Social Relations (impetus: Mosteller offered chair of Statistics at Univ. of Chicago)
- Early recruits include William Cochran (from Johns Hopkins, 1957), John Pratt (from Chicago, 1957), Arthur Dempster (from Toronto, via Bell Labs, 1958)
- Mosteller served as chair 1957-1969, 1973, 1975-1977
- Cochran: fundamental contributions in experimental design, sample surveys, observational studies, influential books *Experimental Design* (1950, with G. Cox), *Sampling Techniques* (1977), *Statistical Methods* (with G. Snedecor)

Pratt, Raiffa, Cochran, Dempster, Mosteller (1959)

Harvard Statistics pre-2000: Long-term faculty

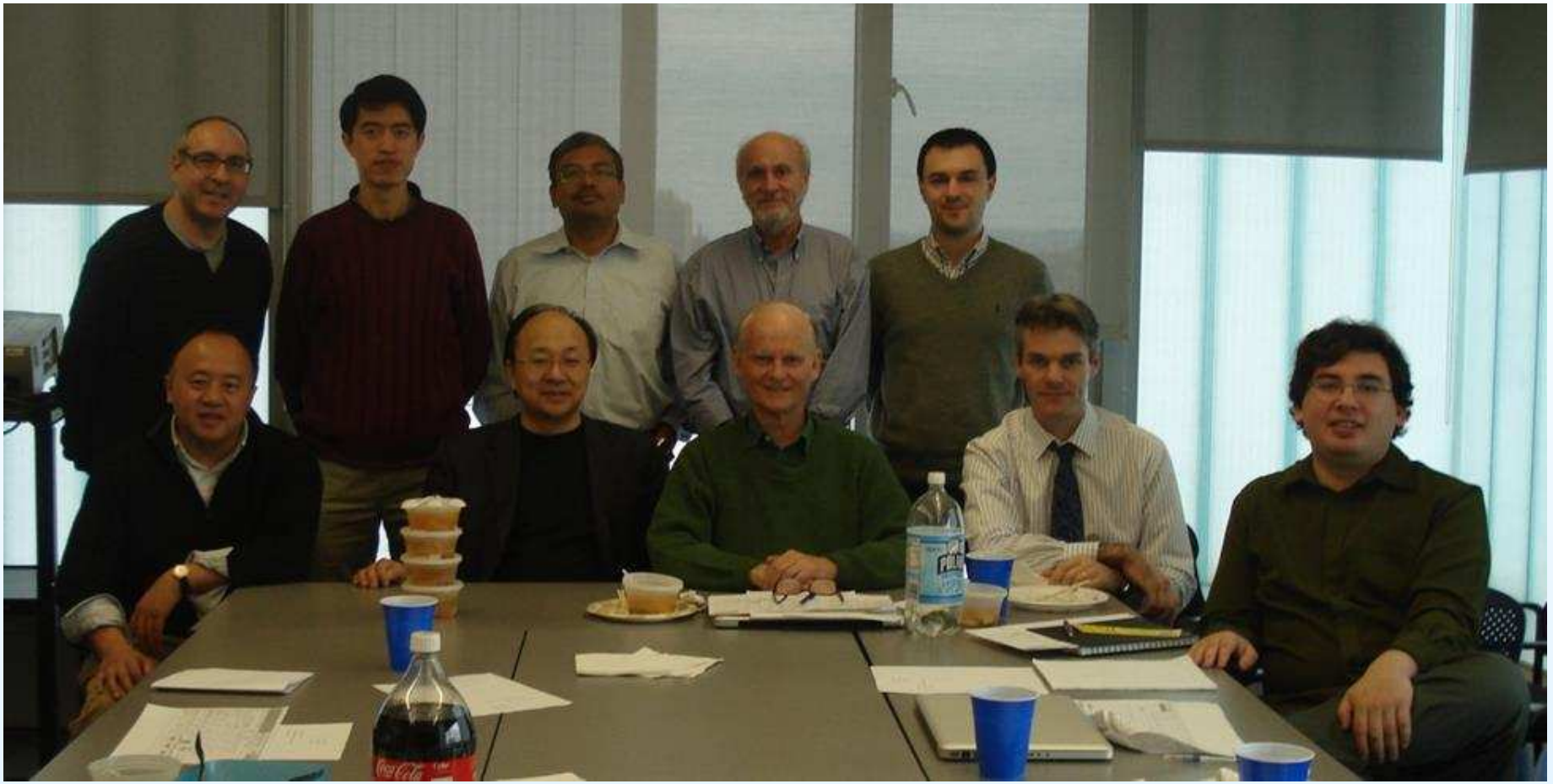
- Art Dempster (1958-) chair 1969-1975, 1977-1979, 1982-1985
  (EM algorithm, Dempster-Shafer theory of belief functions)
- Peter Huber (1978-1988) chair 1979-1982
  (robust statistics)
- Donald Rubin (1984-) chair 1985-1994, 2000-2004
  (EM algorithm, multiple imputation, propensity scores, causal inference)
- Herman Chernoff (1985-)
  (Chernoff bound, Chernoff faces, Chernoff information, Chernoff distribution) Ask about MIT Statistics?
- Carl Morris (1990-) chair 1994-2000
  (empirical Bayes and hierarchical modeling, quadratic exponential families)

Harvard Statistics Department alumni

Harvard Statistics since 2000

- By 2000, only five ladder faculty (three senior).
- Added faculty include Jun Liu (2000-), Sam Kou (2001-), Xiao-Li Meng (2001-), Joe Blitzstein (2006-), Stephen Blyth (2006-), Tirthankar Dasgupta (2008-), Edo Airoldi (2009-), Michael Parzen (2010-), Natesh Pillai (2010-), Luke Bornn (2012-), Luke Miratrix (2012-), Neil Shephard (2013-)
- Xiao-Li Meng chair (2004-2012), helps extend influence of department around university, obtain several new positions
- Dave Harrington, joint with Biostatistics, helps develop undergraduate program and is now acting chair.
- Current rapid growth in undergraduate program, both for undergraduate majors and for popular service courses developed by Joe Blitzstein, aided now by Mike Parzen

Harvard Statistics faculty (2012)

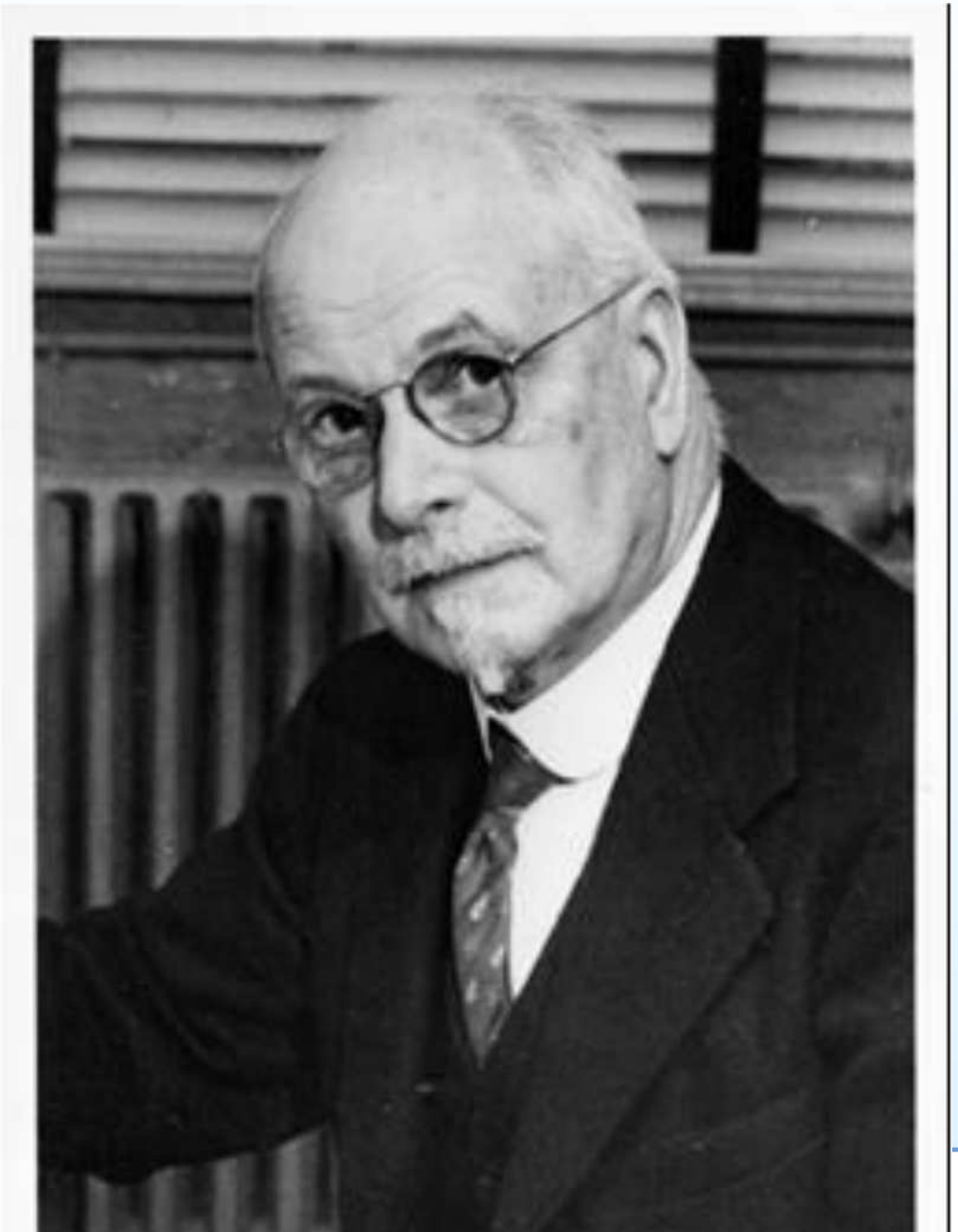


Departmental vignettes: Harvard Biostatistics

(by Nan Laird and Marvin Zelen)

- Harvard School of Public Health founded in 1922, with Vital Statistics one of four core departments. Edwin B. Wilson (who came to HSPH from chairing Physics at MIT) chair from 1922 to 1946.
- Wilson accomplishments include
  - confidence interval for binomial proportion (1927)
  - "Wilson-Hilferty approximation" for chi-squared quantiles (1931)
  - LD50 and its sampling error (with Jane Worcester, 1943)
  - managing editor of *Proceedings of National Academy of Sciences* for its first 50 years (1915-1964)

Harvard Biostatistics chairs: Edwin B. Wilson (1940)



Mosteller influence and transition to Zelen

- Renamed Biostatistics department in 1946, chair Hugo Muench 1946-1961, Robert Reed (1961-1973), Jane Worcester (1973-1977), with department establishing reputation for outstanding teaching and service
- Fred Mosteller (chair 1977-1981) consolidates biostatistics faculty from various departments in medical area, recruits Marvin Zelen from SUNY Buffalo as Professor of Statistical Science
- Zelen founds Department of Biostatistics and Epidemiology at Sydney Farber Cancer Institute (later Dana-Farber) but uses Biostatistics at HSPH for faculty appointments and teaching. With faculty brought from Buffalo (including Colin Begg, Richard Gelber, David Schoenfeld, Ken Stanley, Marcello Pagano, Stephen Lagakos), department triples in size overnight.

Marvin Zelen's decade of leadership (1980-1990)

- Faculty increases from 23 to 34, graduate students from 27 to 60.
- Faculty appointed to Department of Biostatistics at DFCI include James Anderson, Dianne Finkelstein, Rebecca Gelman, Robert Gray, David Harrington, Myrto Lefkopoulou, Louise Ryan, Anastasios Tsiatis.
- DFCI faculty heavily involved in multi-center cancer clinical trials and at Harvard-affiliated hospitals
- Six Cities Study of air pollution and health coordinated by James Ware, with collaborators including Tom Louis, Nan Laird, Victor DeGruttola. Woburn Study by Lagakos and Zelen establishes childhood leukemia cluster, leading to movie "A Civil Action."

Harvard Biostatistics chairs: Laird, Mosteller, Zelen

Harvard Biostatistics retreat (1990)

Harvard Biostatistics since 1990

- Department continues to grow, extend influence under Nan Laird (1990-1999), Stephen Lagakos (1999-2007), Louise Ryan (2007-2009), Victor DeGruttola (2009-)
- New research includes AIDS Clinical Trial Group, group working in Statistical Genetics, Program in Environmental Statistics, partnership with Schering-Plough, Kitasato U.
- New faculty include L.J. Wei, Michael Hughes, Paige Williams, Rebecca Betensky, Tianxi Cai, Christoph Lange, Brent Coull, Xihong Lin, Francesca Domenici, Giovanni Parmigiani, and many faculty in Epidemiology and in Harvard Medical School having secondary appointments in Biostatistics.
- By 2012, department had 80 graduate students, 67 research fellows and associates, 33 research scientists, 58 faculty at primary and secondary level.

Comments, anecdotes?

Thanks!