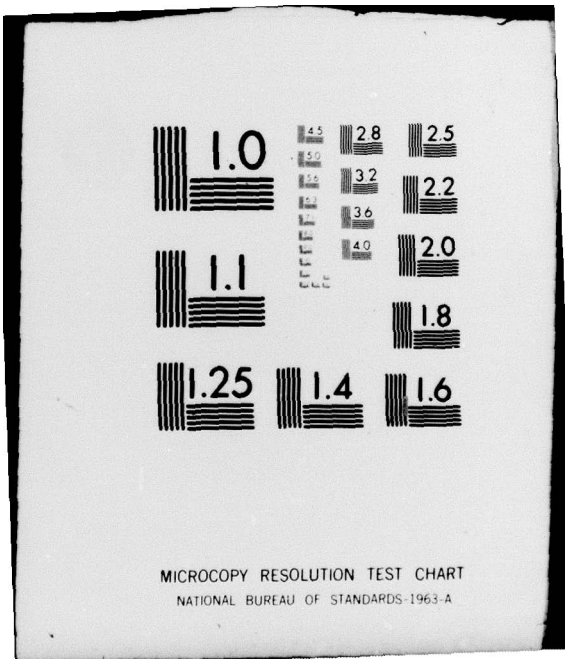
AD-A073 271 TEXAS A AND M UNIV COLLEGE STATION INST OF STATISTICS F/G 5/9
I. DISCUSSION OF STATISTICAL EDUCATION. II. MEMORIAL TO MICHEL --ETC(U)
JUL 79 E PARZEN DAAG29-78-G-0180
UNCLASSIFIED TR-A-8 ARO-16228.8-M NL

| OF |
AD A073271

END
DATE FILMED
9-79
DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

ARO 16228.8-M

TEXAS A&M UNIVERSITY
COLLEGE STATION, TEXAS 77843

LEVEL II

AD A073271

DDC FILE COPY

I, Discussion of Statistical Education,
II. Memorial to Michel Loève .

by Emanuel Parzen

Institute of Statistics, Texas A&M University

Technical Report No. A-8

July 1979

Texas A & M Research Foundation
Project No. 3861

"Maximum Robust Likelihood Estimation and
Non-parametric Statistical Data Modeling"

Sponsored by the U.S. Army Research Office

Professor Emanuel Parzen, Principal Investigator

Approved for public release; distribution unlimited.

DDC
RECEIVED
AUG 30 1979
C

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Technical Report No. A-8	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): I. Discussion of Statistical Education, II. Memorial to Michel Loève	5. TYPE OF REPORT & PERIOD COVERED: Technical
6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s): Emanuel Parzen	8. CONTRACT OR GRANT NUMBER(s): DAAG29-78-G-0180
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Texas A&M University, Institute of Statistics, College Station, TX 77843	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Army Research Office, Research Triangle Park, NC 27709	12. REPORT DATE: July 1979
	13. NUMBER OF PAGES: 11
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): ARO 16228.8-M	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

NA

18. SUPPLEMENTARY NOTES

The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Statistical education, statistical science.
History of probability theory
Reproducing kernel Hilbert spaces.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Part I is a discussion of the report "Preparing Statisticians for Careers in Industry", Committee on Training for Statisticians for Industry of the American Statistical Association. Part II are remarks to be made in memory of Michel Loève at a Memorial Session in Washington, D.C., on Monday, August 13, 1979.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

347 380

79 08 28 029

Discussion of the Report
"Preparing Statisticians for Careers in Industry"
Committee on Training of Statisticians for Industry
American Statistical Association
Section on Statistical Education

by

Emanuel Parzen
Texas A&M University

I welcome the report of the Committee on Training of Statisticians for Industry, since I believe it should be heeded by academic statisticians who are conducting graduate education in the discipline of statistics. The committee's first ten recommendations seem to me to be an excellent summary of the basic requirements of any graduate program appropriate for the 1980's: a balance of theory and practice, course work which includes a wide variety of statistical techniques, hands-on experience with real problems, illustrations in statistical methods courses by real problems and associated data, discussion of the realities of statistical practice, development of skills for communicating about statistical reasoning with non-statisticians, extensive familiarity with statistical computing, opportunities for students' research which are stimulated by real problems, and faculty members genuinely interested in real problems. It seems to me unwise to attempt to set goals for graduate education oriented towards training statisticians specifically for industry. The beauty of

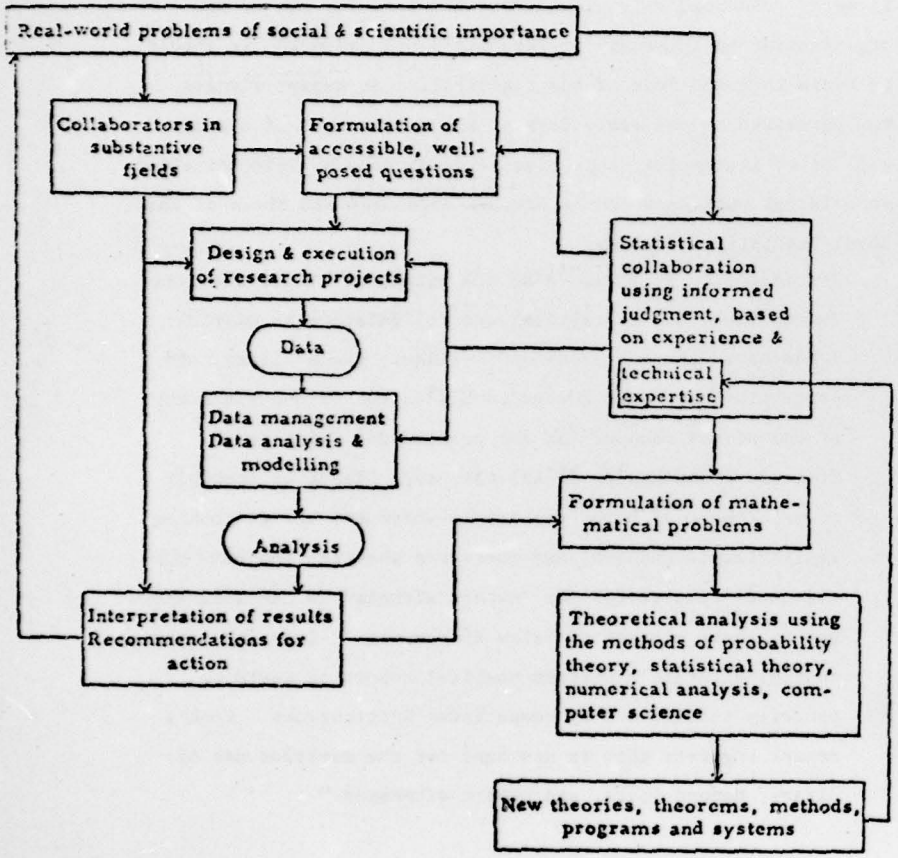
the discipline of statistics is that it enables an individual to have a multitude of careers, and individual statisticians should be educated so that they are able to alternate in their careers between industry, government, and universities (the profit, public, and non-profit sectors of the economy). The committee's last four recommendations should be viewed as means for accomplishing this goal.

At the 1978 statistical meetings in San Diego there was much discussion of the need to educate (not train) a "whole statistician" in one body (see discussion by Tukey to Parzen (1979)). It seems to me that the most effective way to implement the committee's recommendations is to develop graduate programs in statistics which educate "statistical scientists" able to conduct collaborative research; their roles can be depicted in a flow chart (Figure A). A point emphasized by this flow chart is that the statistician working on applications must not only deliver statistical care to important scientific and social problems, but must also be oriented to providing feedback from his problem-solving livelihood activities to the development of the overall discipline of statistical science. This remark seems to me to explain <u>how</u>, and <u>why</u>, statistical consulting should be conducted as part of graduate programs in statistics.

The problems involved in designing educational programs for statisticians seem to me to be magnificiently exemplified by an old Chinese saying ascribed to Kuan-tzu:

Figure A

Flow Chart Depicting the Roles of a Statistical Scientist



"If you give a man a fish,
He will have a single meal
If you teach him how to fish
He will eat all his life".

I interpret this proverb to mean that one can distinguish between two types of "statistical methods" courses, which I call fish courses and fishing courses, to which category a particular course belongs seems to depend on the background of the statistician making the judgement. I define a fishing course to be one that encourages the student to think creatively and imaginatively about his or her scientific research problems and the role of modern theories of statistical data analysis and modelling. A fish course usually describes without much proof the mathematical solutions of various routine statistical analysis problems based on probability models for observed data whose validity is not usually checked. The significant feature that I believe will make the discipline of statistics very different in the next decade from what it has been in the past is our increasing ability to test the goodness of fit of models.

An important guideline for universities to consider in developing programs for educating statisticians who will work in industry is that this be but one of the dimensions in a broad program which (1) emphasizes that the discipline of statistics is currently experiencing a healthy agitation concerning its theories, methods, and interactions with other disciplines, and (2) encourages faculty

members to be not locked into the ways of thinking that they learned in the particular graduate schools they attended. I would conjecture that what industry wants to hire is a graduate who has been "turned on" by a challenging education which has provided him or her in many courses with the experience that statistical thinking can (1) be beautiful, (2) work to provide scientific insight from data, and (3) provide conclusions and recommendations which can be communicated to a concerned audience.

In discussing the education of practicing statisticians (or statistical scientists) I believe it important that we be concerned with the development of public recognition of the professional identity of statisticians. Professor H. O. Hartley in his 1979 Presidential Address entitled "Statistics as a Science and as a Profession" calls attention to the problem of researchers trying to act as their own statisticians. I strongly endorse his views concerning the answer to this problem. "Our answer should be one of strength, namely to practice first rate statistics, statistics which is of a quality that our offer of help is an offer that the other fellow cannot refuse! Cannot refuse because a refusal would waste his time doing dilatantic statistics instead of bringing his sound expertise to bear on the fruitful cooperation with the professional statistician who knows what he is doing."

Any program for the education of practicing statisticians should include a discussion of the question: in explaining to clients the role of the statistician, is it necessary to caution that the statistical scientist should not be primarily an expert witness, concerned only with making the case for his or her organization in adversary or regulatory proceedings? We should be aware that the fear of the statistician as expert witness was perceived in the early days of the development of the discipline of statistics; this discovery was made by Fred Mosteller in a letter published in the October 1978 News and Notes of the Royal Statistical Society:

> "Professor D. E. Barton asks the origin of "There are lies, damned lies, and statistics" and its relation to Disraeli. I had heard the quotation as "---liars, damned liars, and statisticians." Sir Edward Cook in a two-volume biography of one of the founders of our profession, The Life of Florence Nightingale, pp 433-434, says "There is probably no department of human inquiry in which the art of cooking statistics is unknown, and there are skeptics who have substituted "statistics" for "expert witness" in the well-known saying about classes of false statements." Cook was doubting that Nightingale's uniform hospital reporting would be properly carried out by competitive institutions. Cook's remark suggests that we now hunt for the earliest use of "liars, damned liars, and expert witnesses."

My final comment relates to the question: are there any effective steps available to the committee to make universities pay attention to its recommendations? Somewhat facetiously, I would like to point out that an effective step would be the publication of an "Annual Consumer's Report for Students Seeking Graduate Education in Statistics" which would provide up-to-date information on how universities rate by the reasonable criteria that the committee is recommending. Such a report would also be useful to universities to convince their administrations to give them the resources required.

Reference

Parzen, E. (1979) "Nonparametric Statistical Data Modeling" Journal of the American Statistical Assn., 74, 105-131.

In Memory of Michel Loève

by

Emanuel Parzen

Texas A&M University

When we pause here to remember our colleague, Michel Loève, it seems to me to be an occasion when those of us who had the good fortune to be among his students can communicate to you here some of the spirit that he taught us so that you can count him among your teachers. Thus I speak to you as one who was a graduate student at Berkeley from 1949 to 1953, and as Loève's first Ph.D. student.

Loève was born on January 22, 1907 in the Middle East in an area whose correct geographical description seems confusing to many. It seems to me historically incorrect to say that he was born in Palestine or in Israel, neither of which existed at the time. Loève was born in the Land of Israel (Eretz Israel) when it was part of the Ottoman Empire. Loève came to France in the early 1930's and in 1941 received his doctor of science degree in probability theory from the Sorbonne. It should also be noted that he became an Actuary in 1936. During the Second World War he was interned in the Drancy concentration camp, his life saved by the fact that he had an Egyptian passport. Not being a French citizen, he could not receive a university position in France, and in 1948 he became Professor at Berkeley at the invitation of Jerzy Neyman. He died this year, at the age of 72, on February 17, 1979.

Footnote: Remarks for a session dedicated to the memory of Michel Loève to be held at the Institute of Mathematical Statistics meeting in Washington, August 1979.

Loève is world famous for his treatise on probability theory, first published in 1955, and now in its fourth edition. The research for which he is most remembered is undoubtedly that on random functions of second order (the foundations of time series analysis) which appears in an important appendix written by Loève in Lévy's 1948 book on Brownian motion. Loève's name is immortalized in the Karhunen-Loève expansion in eigenfunctions and the Cramer-Loève spectral representation.

Loève's research also included contributions to the probabilities of systems of events, limit theorems for laws of sums of dependent random variables, almost sure convergence, and ranking limit theorems.

Loève was an inspiring teacher in every aspect of being a teacher. He tried to teach us to be cultured and to cultivate our roots. Roots is a modern word used to denote one's inherited culture, but Loève taught us to know also our mathematical roots.

Loève was the only living doctoral student of Paul Lévy; the only other student of Lévy was W. Doeblin whose name should be well known to every researcher in Markov processes. (It is appropriate on this occasion to remember also Doeblin who was a German Jewish refugee who served in the French Army and never returned.) When I received my Ph.D. under Loève, I became Lévy's grand-student (and incidentally a great grand-student of Hadamard). All my Ph.D. students are therefore Lévy's great grand-students. The Ph.D. students of my Ph.D. students are my grand-students, and thus the geneology grows.

An important lesson I learned from Loève is the BROTHERHOOD OF MATHEMATICIANS. Loève told me the story when I was his student that when he came to Paris as a student, he introduced himself to Hadamard who interceded with the bureaucracy to obtain for Loève various student privileges

at the University of Paris. Said Loève to Hadamard, "Why do you help me so much when you hardly know me?" Responded Hadamard: "I have no choice. When you ask my help in order to become a mathematician, I must help you -- all mathematicians are brothers."

Finally, I believe there is an interesting and informative tale to be told about REPRODUCING KERNEL HILBERT SPACES (RKHS) which illustrates the social processes by which mathematical facts and proofs are propagated. It seems to be fair to say that RKHS are regarded by mathematicians as relatively advanced tools; yet RKHS are accepted in the statistical and engineering literature of time series analysis and statistical communication theory, and used as efficient techniques for obtaining insight into solutions of problems of statistical inference. My research from 1957 to 1961 was the first to demonstrate these facts. It is interesting to trace how I happened to know about the then brand new field of RKHS, since Aronszajn's basic paper appeared only in 1951. I studied RKHS extensively in 1951 as my Ph.D. prelim topic, proposed by Loève who knew about RKHS from his fellow student Aronszajn whose doctoral thesis on RKHS was also Paris 1941. The important theorem on the isomorphism between Gaussian processes and RKHS was given by Loève in his appendix to Lévy's 1948 book.

Other students of Loève could tell you of other things that they learned from him. Through his books, his students, his students' students, and so on, each of you have undoubtedly been students of Michel Loève. May his name be long remembered, as long as probabilists assemble.