

# Training Majors & Graduate Students Beyond Mathematics

Sastry G. Pantula

North Carolina State University

*Training Problem Solvers*

# Outline



- Overview of our Department
- Undergraduate Program
- Graduate Training
- Departmental Culture
- NISS/SAMSI
- Interdisciplinary Programs
- Conclusions

# Overview



- One of the "First" Statistics Departments in the country
- Gertrude Cox in 1941- first female Statistician to Head a Statistics Dept.
- Strong foundation in Mathematical Statistics, with an eye towards applications
- Train Problem Solvers for the future

# Overview



- Diverse faculty cover most topics in Statistics
  - Biomedical Statistics
  - Environmental Statistics
  - Industrial Statistics
  - Mathematical Statistics
  - Statistical Genetics
  - Time series/Econometrics
  - Statistical Computing

# Overview



## ■ Diverse Faculty

- 8 Assistant Professors
- 8 Associate Professors
- 18 Full Professors (One is a Dean)
- 4 Teaching Assistant Professors (5-year)
- 3 Research Assistant Professors
- 3 Post doctoral fellows
- 2 Research Scientists

(Math has 60 faculty; Good interactions)

# Overview



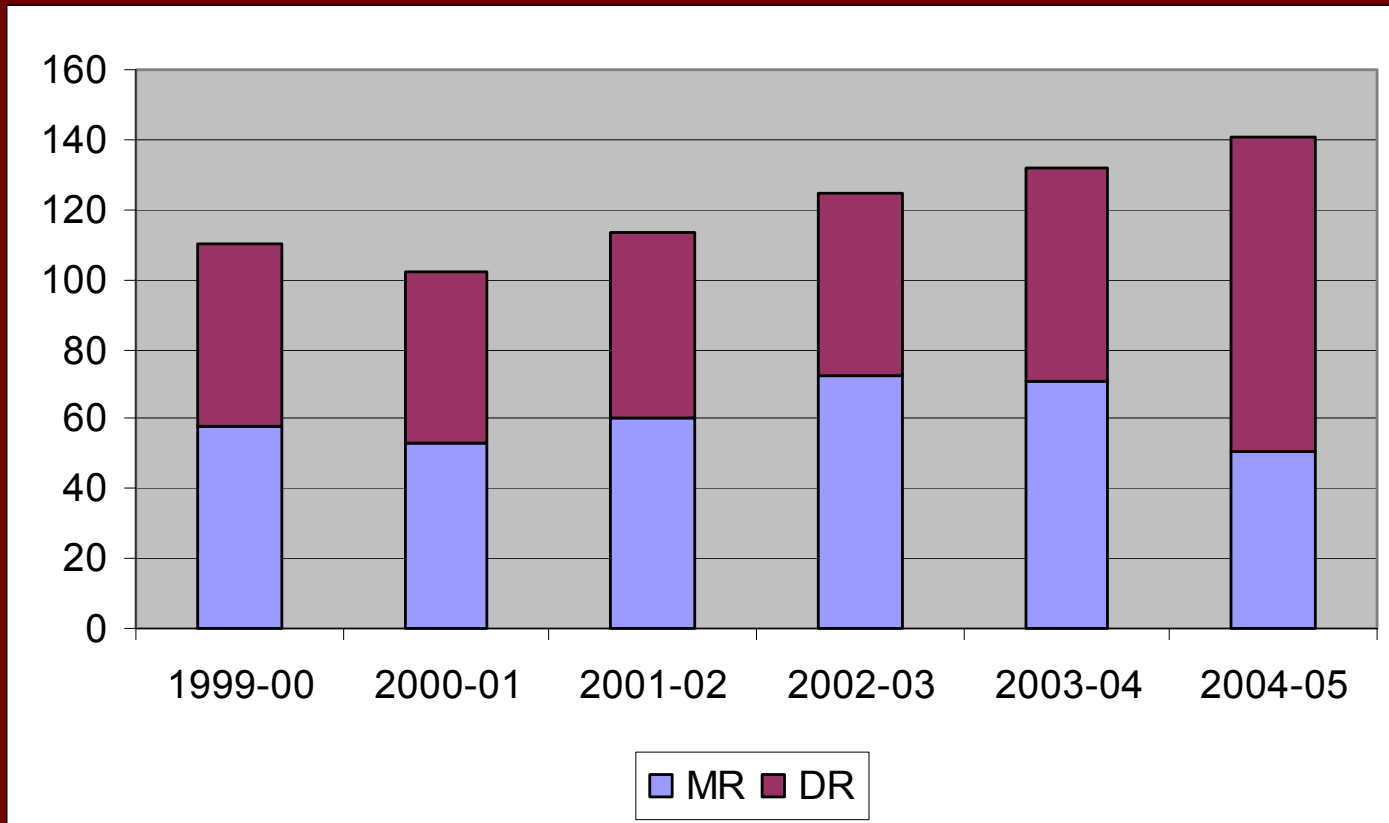
- Diverse Faculty
  - 13 females
    - 1 Hispanic
    - 3 African
    - 4 Indian
    - 6 Chinese
    - 32 White



# Overview

- B.S. In Statistics
- Master of Statistics (and M.Sc.)
- Ph.D. in Statistics
- Interdisciplinary Graduate Degrees
  - Operations Research (M.S and Ph.D.)
  - Biomathematics (MS and Ph.D.)
  - Bioinformatics (MS and Ph.D.)
  - Financial Mathematics (MS only)

# Graduate Enrollment





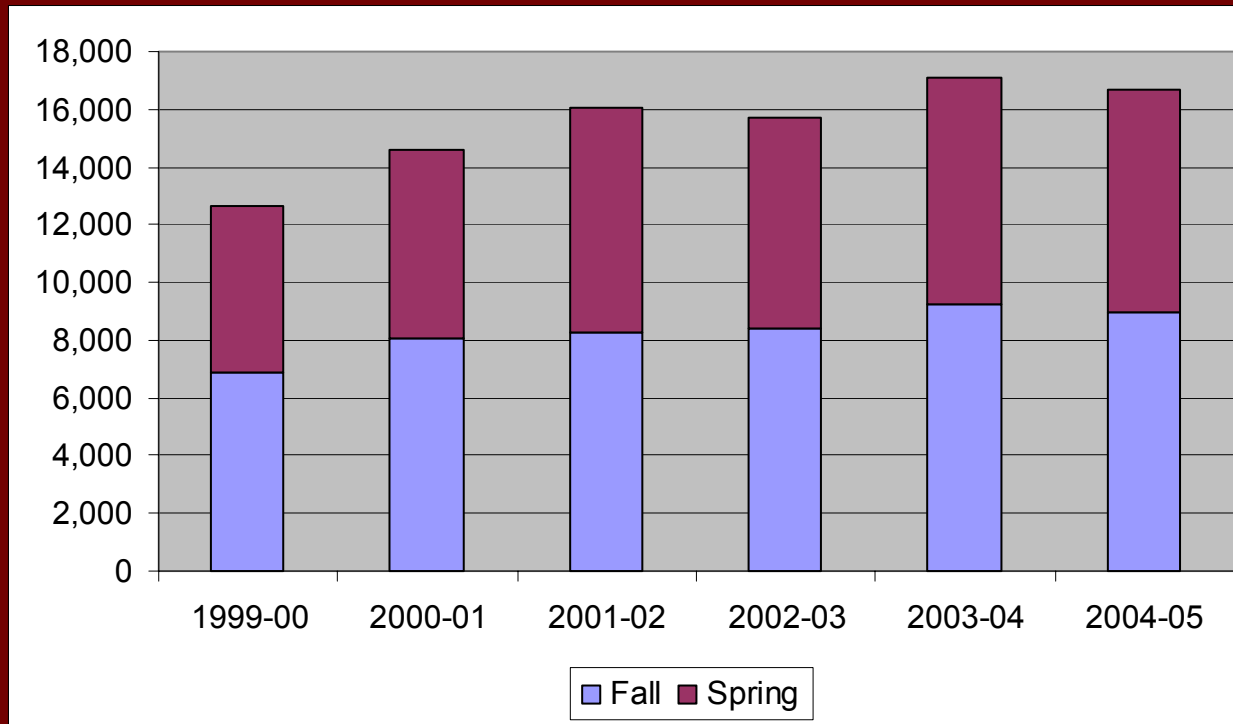
# Graduate Enrollment



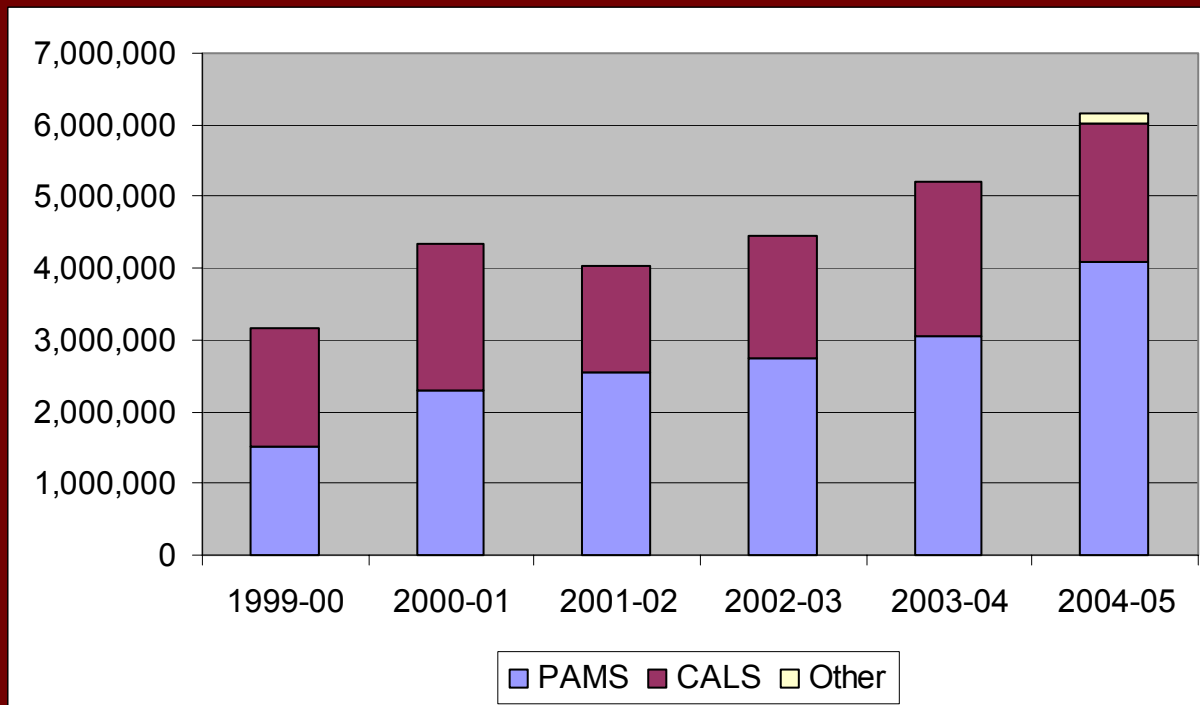
- Fall 2005
  - 38 Graduate Students
  - 22 US
    - 16 Phd- 9 Female
    - 6 MS – 3 Female

16 International – 9 Female  
- 13 Phd

# Credit Hours



# New Grants Awarded



# Undergraduate Program



- Strong Undergraduate Program
- Second largest in the country
- 90 undergraduates at present
- Graduate typically 20-30 a year
- About 40% go to Graduate School (25%)

# Undergraduate Program



## ■ Recruiting

- Visit High Schools
- Be involved with mentoring- “Girls on Track”
- AP Statistics
- WISE Program
- “Statistics by Example”
  - Invite students with high SAT scores on campus
  - Daily News Paper articles; *Chance* Magazine
  - You can only check in, but, can’t check out!

# Undergraduate Program



## ■ Retention

- Peer Mentoring
- VIGRE Mentoring (grad/post-doc trainees)
- Professional Mentoring
  - Provide experienced and knowledgeable role models for students seeking careers in Statistics.
  - Provide informal advice and counsel for students seeking to clarify their career directions.
  - Increase students' awareness of the variety and scope of career opportunities to Statisticians in NC.

# Undergraduate Program



## ■ Curriculum

- Applied Courses
- Math Courses (Calculus/Linear Algebra)
- Mathematical Statistics Courses
- Applied Courses!
- Statistical Computing
- Research Experience
- Train to be a Problem Solver
- Stat Club trip to DC

# Other Majors



- Encourage students to take
  - a second or third course in statistics
  - both an applied course and a theoretical course
  - a course on simulation techniques (integrate it in all follow up courses)
  - Research practicum courses
  - An opportunity to tutor others  
(We get a large number of transfer students)



# Research Examples

- Joseph McMichael and Ronnie DeFrancis, "Season - Long Perspective on Ozone/Precursors Interaction Among Seven Sites Near Nashville, TN".
- Daric Harrington, "Forecasting Ozone with Yesterday's Data and Meteorological Data in the Charlotte Metropolitan Area".
- Jay H. Bibby and Kimberly A. Madsen, "Is there a relationship between water chemistry and bug abundance?"



# Research Examples

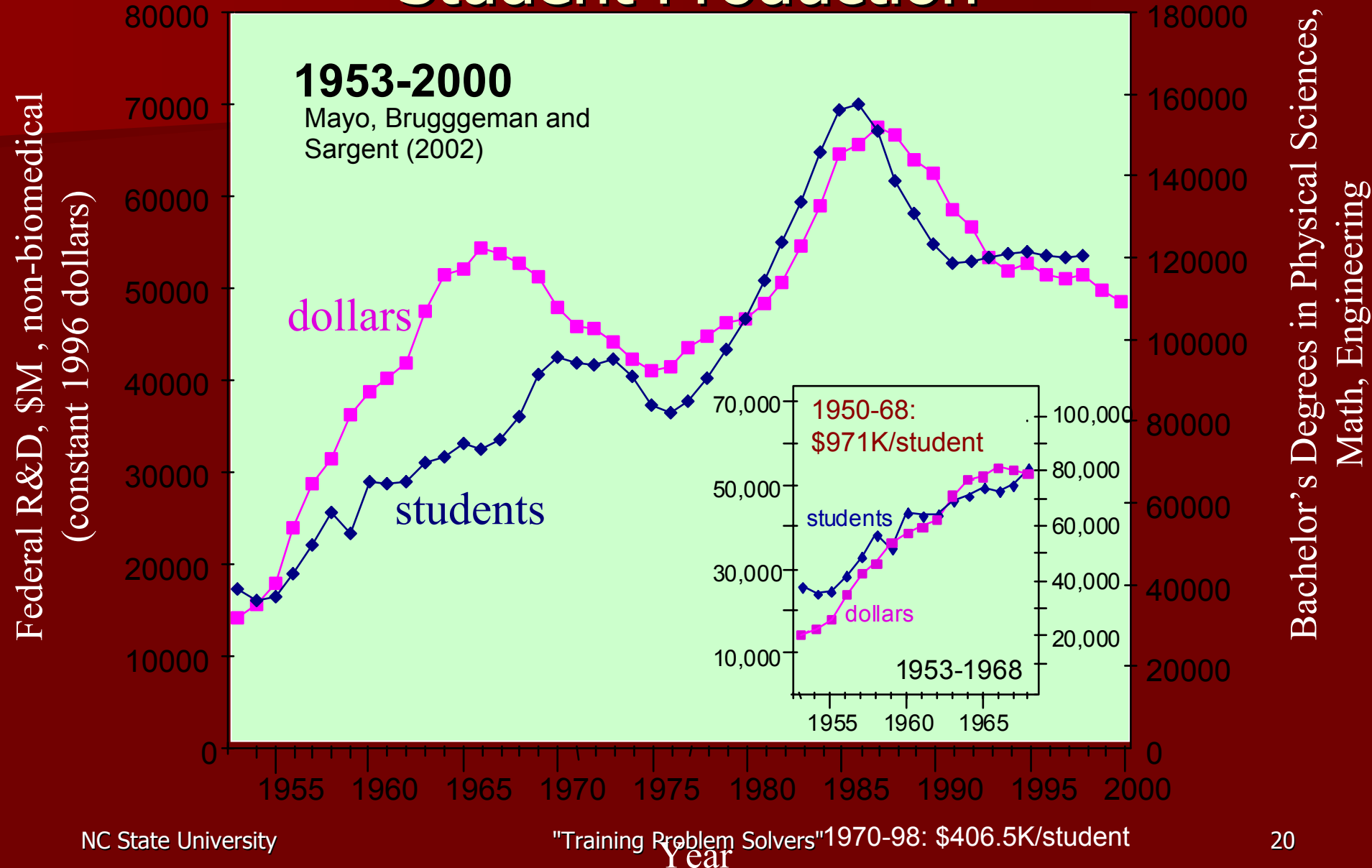
- Jason Grissom worked on a project for the US Department of State. Jason co-authored a report entitled, **Comparison of Particulate Matter Levels in Worldwide Megacities**, report prepared for Ms. Claire Huson and Mr. Ken Doolan, U. S. Department of State, August 17, 2000. Jason briefed State Department Officials on the results of his work. **(Study cited in USA Today's All-Star Academic Teams. Jason made the second team).**

# Research Examples



- W. Hunt, N. Shah, K. Weems, M. Stephens, W. Swallow, M. Crotty and A. Moore, **"Training Environmental Statisticians – Tomorrow's Problem Solvers,"** The First United States Conference on Teaching Statistics (USCOTS), Ohio State University in Columbus, Ohio, May 19-21, 2005.

# Federal R&D Funding Dictates B.S. Student Production



# Graduate Programs



## ■ Recruiting

- Over 50% of our graduate students are math majors; 20% stat majors;
- Joint grant with Spelman- Infinite Possibilities
- Visit St. Olaf, Meredith College, and others
- Be active in MathFest, StatFest and SACNAS
- Visit REUs like SUMSRI
- Have VIGRE REUs
- SIBS

# SIBS



- "How I became interested in statistics:"
- Personal histories of some SIBS instructors
- Randomized trials vs. observational studies
- Descriptive statistics and probability
- Introduction to SAS statistical software
- Clinical Research and the PURSUIT Trial (an example of a large clinical trial, discussed by the lead physician)
- Population, sample, and sampling distribution
- Random sampling on the computer
- Hypothesis testing, estimation, and confidence intervals
- Field trip to DCRI (Lecture by Dr. Robert Califf, Director of DCRI, on the importance of biostatistics at DCRI)
- Durham Bulls Baseball game

# SIBS



- Introduction to clinical trials: Randomization, sample size, conduct of trials, and the intent-to-treat principle
- Designing a clinical trial and planning the statistical analysis
- From lab to patient: Statistics in drug development
- Field trip to SAS Institute
- Trip to Exploris Museum and IMAX theater
- Associations among variables and introduction to linear regression
- Using statistical software for estimating associations and regression
- Introduction to multiple linear regression
- Using statistical software for multiple linear regression analysis
- Introduction to logistic regression analysis
- Using statistical software for logistic regression analysis
- What causes heart disease? Introduction to observational studies and confounding
- Field trip to GlaxoSmithKline

# SAMSI Help



- Two-day Undergraduate Workshops on weekends during the academic year
  - For juniors and seniors
  - Description of SAMSI programs, and hands-on experience with one of them
- SAMSI/CRSC Undergraduate Workshop for one week in early June
  - For juniors, seniors, and high-school teachers
  - Description of SAMSI programs and intensive hands-on experience in use of statistical and mathematical models to analyze experimental data collected in the CRSC/Math Instructional Research Lab



# Graduate Programs



## ■ Curriculum

- Strong Masters Program (Not a consolation Prize)
- MS Program with Concentrations
- Statistical Consulting Required
- Good mixture of Applied, Theoretical and Computational Courses
- 5-year BS-MS program

# Graduate Programs



- Doctoral Program in Statistics
  - Solid foundation in mathematics
  - Most theses are motivated by a practical problem
    - **Chen, Li.** Bayesian Hierarchical Spatial-temporal Models for Wind Prediction.
    - **Gosky, Ross Matthew.** Bayesian Analysis and Matching Errors in Closed Population Capture Recapture Models.
    - **Hwang, SangPil.** Time Series Analysis using Logistic Function.
    - **Lokhnygina, Y.** Topics in Design and Analysis of Clinical Trials.
    - **Lu, Na.** Statistical Issues in Coherent Risk Management.
    - **Remlinger, Katja Sabine.** Statistical Design and Analysis of High Throughput Screening Data using Pooling Experiments and Data Mining Techniques.

# Graduate Programs



## ■ Ph.D. Co-major Option

- Economics
- Computer Science
- Genetics

Examples:

**Chao, Hui-Ping.** Modeling nonlinearity in US livestock cycles.

**Chun, Heuiju.** Probabilistic and statistical modeling for geometric structure of nonwoven fabrics

**Hilmer, Christiana Ruefli.** A comparison of resampling techniques when parameters are on a boundary: the bootstrap, subsample bootstrap, and subsample jackknife.

# Broadening



## ■ Courses

- Consulting Class is required
- “Preparing for Research”
- “Statistics in Pharmaceutical Research”
- Data mining courses
- Wide choice of courses
- Fellowships require additional courses
  - SAS (Computing/Numerical Analysis)
  - Merck and Lilly (Chemistry/Biostatistics courses)

# Broadening



## ■ Training

- Preparing the Professoriate Program
- Stat-Teach Group; “Teaching Intro Stat”
- Industrial Mathematics and Statistical Modeling Workshop for Graduate Students
  - 9 days: end of July each year
  - 6 teams work on “industrial” projects presented by experienced scientists and engineers: 6 students, 1 problem owner, 1-2 faculty mentors

# Broadening



- Mentoring Lunches, esp. for students from HBCU and smaller colleges
- Invite folks from Industry and Academia
  - Role models
- Graduate Industrial Traineeship Program (Gerig, 2005 JSM Presentation)

# Graduate Industrial Trainee (GIT) Program

- Program for graduate students
- Mentored experience in statistical practice
- Trainees remain full time students while working on-site at participating industries
- Provides funding for graduate students
- Promotes interaction between faculty and industry scientists

# Supervision and Mentoring

- Each GIT has a campus advisor and a GIT supervisor/mentor
- Supervisor/mentors are almost always statisticians
- Some are alumni and Adjunct faculty
- Campus advisors and GIT supervisor/mentors keep in touch



## The Partner's Perspective

- GIT program is not a source of cheap labor but trainees do contribute
- GIT program is a way for industry to support and participate in the recruitment and training of future employees (SPAIG)
- Host units often hire trainees after they complete their degrees

GlaxoSmithKline	SAS Institute
Research Triangle Institute	Becton Dickinson
Burroughs-Wellcome	Nortel
CIIT	EPA
PPD Pharmaco	Triangle Pharmaceuticals Inc.

Analytical Sciences Inc.	Constella Group
Quintiles Inc.	Inspire Pharmaceuticals, Inc.
Clintrials Inc.	Raleigh News & Observer
Inveresk	Cuddy Farms
USGS	USDA Forest Sciences Lab

# Broadening



- Industry – University Relationships (Starbuck, 2005 JSM)
  - NSF Funding (GOALI)
  - Invite industry/government folks for seminars
  - Include an industry rep on review teams
  - Encourage sabbatic leaves outside academia
  - Encourage students and postdocs to work on some applied problems

# Broadening



- Make use of Institutes (SAMSI)  
(Berger, 2005, JSM)

# More Details on 2005-06 Programs

- Financial Mathematics, Statistics and Econometrics (FMSE)
- National Defense and Homeland Security (NDHS)
- Astrostatistics (ASTRO)

# FMSE (Fall 2005)

- Focus on bringing together applied mathematics, economics, finance, statistics and econometrics
  - Address essential tasks of modeling, incorporating data, and computing in domains ranging from financial and energy derivatives to real options
- *Program leaders:* Marco Avellaneda (NYU), Jean-Pierre Fouque (NCSU; co-chair), Eric Ghysels (UNC; co-chair), Ronnie Sircar (Princeton), Ruey Tsay (Chicago), Thaleia Zariphopoulou (Texas); John Lehozcky (CMU; NAC liaison); Ralph Smith (directorale liaison)

# NDHS (full year)

- Possible working groups
  - Biointelligence (with connections to CDC Biosense)
  - Anomaly detection
  - Data integration, confidentiality, quality
  - Real time inference (data streams)
  - Dynamics of massive databases
  - Social networks
- *Program leaders:* David Banks (Duke, LDC liaison), James Crowley (SIAM), Lawrence Cox (NCHS; co-chair), Jon Kettenring (Drew University), Nell Sedransk (NISS; co-chair); Sallie Keller-McNulty (LANL; NAC liaison);

# ASTRO (Spring 2006)

- Potential emphases (potential synergy with NDHS)
  - Exoplanets
  - Handling selection bias in astronomical surveys
  - Anomaly detection in massive data sets
  - Summarizing statistical information for future science
  - Source detection
  - Adaptive exploration
  - Physics/statistics interfaces
- *Program leaders:* Jogesh Babu (Penn State; chair), Alanna Connors (New Hampshire), Eric Feigelson (Penn State), Donald Richards (Penn State), Larry Wasserman (CMU); Peter Bickel



# High Dimensional Inference and Random Matrices (Fall 2006)

- Potential emphases
  - Extreme sample eigenvalues
  - Properties of sample eigenvectors
  - Empirical distribution of eigenvalues
  - Design of snapshots for computer model approximation
  - Nonlinear/topological approaches to dimensional reduction
  - Bayesian utilizations of random matrices
  - Stochastic evolution of random matrices
  - Statistical issues involving EOFs in climatology
- *Program leaders:* Iain M. Johnstone (Stanford; chair), Craig A. Tracy (UC Irvine), Ken McLaughlin (Arizona); Peter Bickel (NAC liaison); Chris Jones

# Development Assessment and Utilization of Complex Computer Models (full year)

- Three subprograms, with complementary emphases, will be chosen from the domains
  - Engineering modeling
  - Environmental / ecological modeling
  - Biomedical modeling
  - Social network modeling
  - Microsimulation modeling
- Program Leaders: Peter Reichert (ETH), Tom Santner (Ohio State), Henry Wynn (LSE)

# Mentoring



- Peer mentoring
- Mentoring by post-docs
- Mentoring lunches
- Faculty mentoring
  - Reported on annual reports
  - Reviewed
  - Rewarded (assuming that we get raises)

# Interdisciplinary Work



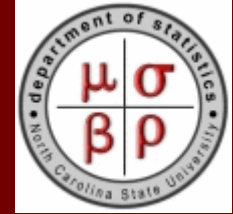
- Training at the intersection of biological and mathematical sciences
- Communicate across fields
- Breadth vs Depth
- Core vs Interdisciplinary
- Downtown vs Suburban Sprawl
- Excite them to venture into other areas

# Other Successes



- Our Math Department has an Excellent Center for Scientific Computing
- FRG in Financial Mathematics
- St. Olaf's MCTP program
- VIGRE programs at various places
- Biomathematics at Utah
- REU at Nebraska

# Summary



- Strong foundation is important
- Invest in Broadening
- Collaborate with other scientists
- Mentor!
- Integrate Research into Education
- Take advantage of folks in Industry/Govt
- Train Critical Thinkers, Problem Solvers and Decision Makers!

# Can't Afford to Give Up!

- Need to continue to make every effort to recruit, retain, mentor and graduate!
- Survey indicates High School students are not challenged enough in math
- Tap into students that were interested in computer science- bioinformatics, computational biology, scientific computing

# Final Pitch

- Diversity – look for opportunities to work with HBCU's and Women's Colleges
- Faculty at these colleges are key to our pipeline
- Share your expectations with them
- Hire women and minorities as faculty members (on-going Diversity Committee)
- Broadening in terms of Diversity!



# Salary Survey

- [http://www.amstat.org/profession/salaryreport\\_acad2004-5.pdf](http://www.amstat.org/profession/salaryreport_acad2004-5.pdf)
  - Salary survey for faculty
- [http://www.amstat.org/profession/SPAIGs\\_salarysurvey03.pdf](http://www.amstat.org/profession/SPAIGs_salarysurvey03.pdf)
  - Salary survey for BS, MS and Phd's in industry and government

**Thank You!**