Annual Report
2015-2016
The Department of Statistics and Data Sciences is an academic unit housed in the College of Natural Sciences that supports the statistical and data science needs of The University of Texas at Austin campus.

Our Mission

Our mission is to be a world-class center for statistical science that advances scholarship across The University of Texas at Austin and prepares graduates to flourish in an information-rich world.

Our Vision

The Department of Statistics and Data Sciences will become a vibrant community of scholars that has a strong intellectual identity and is widely recognized for excellence in interdisciplinary research and high impact educational programs.

Our Core Values

- Commitment to education
- Excellence in research
- Broad engagement with science and society
- Collegiality and integrity
From the Chair

In May 2017, the Department of Statistics and Data Sciences (SDS) will celebrate the 10-year anniversary of the annual UT Summer Statistics Institute (SSI). An anniversary (and an annual report) offers an opportunity to reflect upon the past. In 2007, when the first SSI was imagined and brought to fruition, the idea that a department of statistics would be a reality within ten years would have seemed like wishful thinking. Just as SSI has grown by leaps and bounds over the course of a decade, so has the statistics hub within the College of Natural Sciences, from the founding of the Division of Statistics + Scientific Computation to its transformation into the Department of Statistics and Data Sciences. This growth is due to through the sheer will and determination of a small group of dedicated faculty, staff, and students with the essential support of the CNS Dean’s office as well as the upper administration.

Although this annual report includes only a quick look at the many exciting activities in the Department of Statistics and Data Sciences over the past year, I hope that it conveys to you at least some of the quality, impact, innovation, and potential of the department.

One of the highlights of the year was the announcement that we, along with collaborators in Biology and Computer Science, were awarded a $1.1M interdisciplinary NIH T32 predoctoral training grant in biomedical big data sciences. Not only will this prestigious award help us with recruiting the best students for the new PhD in Statistics program but this grant will also support the training of five trainees per year for the next five years.

I am delighted to note the recognition of two of our outstanding faculty. On the international stage, Dr. James Scott was recognized as one of the leading Bayesian statisticians under 35 years old by the International Society for Bayesian Analysis (ISBA). As the recipient of the Susie Bayarri Award, James gave a plenary lecture, “Empirical Bayes and penalized likelihood,” during the world meeting of ISBA. On the local stage, Dr. Kristin Harvey was awarded a 2016 Teaching Excellence Award by the College of Natural Sciences for her exceptional efforts teaching and transforming the curriculum for one of our most highly subscribed service courses, SDS 302 “Data Analysis for the Health Sciences.”

In fact, Dr. Harvey spearheaded an initiative of which we are very proud—the addition of a final research project and poster presentation in the SDS 302 Data Analysis for Health Sciences and SDS 328M Biostatistics courses. This year approximately 2500 undergraduate students designed and completed real research projects, culminating with a poster presentation. This experiential, student-centered activity provided students the opportunity to engage in analyzing real (and often messy) data through an inquiry—based process. Through this project, students developed competencies in the field while being challenged to think creatively and independently.
The faculty have active research programs funded by NIH, NSF, as well as other federal agencies. These funded projects include both methodological and collaborative research. Some examples of the former include development of new approaches for missing data and causal inference, development of new methods for tumor heterogeneity using Next-Generation sequencing data, and development of theory and algorithms for conducting probabilistic inference for Big Data. Some examples of the latter include projects involving determining biomarkers for Duchenne’s muscular dystrophy and assessing the response of hospitals to Medicare readmission penalties.

In other news, faculty continue to serve on the editorial boards of the top journals in statistics and biostatistics both as editors and associate editors (including the Journal of the American Statistical Association, Annals of Statistics, and Biometrics) and publish extensively in them. Faculty also serve on national and international committees for the International Biometrics Society, ISBA, and the National Academies of Science.

I am pleased to report a milestone reached for the PhD in Statistics program—our first three students entered candidacy this year. This is an important step in the life-long journey of learning and scholarship for the candidates, and a notable marker of progress in the program’s mission to prepare of high-quality researchers who will serve as leaders in the field of statistics and data sciences.

I would also like to take this opportunity to welcome Maurie McInnis to UT-Austin as the new Executive Vice President and Provost, and to thank President Gregory Fenves and CNS Dean Linda Hicke for their continued support.

Thank you for another great year.

Dr. Mike Daniels
Table of Contents

I. SDS Highlights in 2015–16 .................................................................................................................... 1

II. Unit Reports

 A. Instruction ............................................................................................................................................. 5
 B. Consulting ............................................................................................................................................... 9
 C. Graduate Degree Programs .................................................................................................................. 11
 D. Graduate Fellows Program .................................................................................................................... 13
 E. Portfolio and Certificate Programs ....................................................................................................... 14
 F. SDS Seminar & Statistics in Action ...................................................................................................... 17
 G. Summer Statistics Institute .................................................................................................................. 18
 H. Corporate Sponsorships ....................................................................................................................... 19
 I. Grants .................................................................................................................................................. 20
 J. Development ........................................................................................................................................... 22

III. Appendices

 A. Organizational Chart .......................................................................................................................... 23
 B. SDS Core Faculty ............................................................................................................................... 24
 C. SDS Seminar Series Speakers ........................................................................................................... 29
 D. Statistics in Action Attendance ........................................................................................................ 30
I. HIGHLIGHTS IN 2015–16

NIH T32 Predoctoral Training Grant in Biomedical Big Data Sciences
In March 2016, the NIH awarded Dr. Michael Daniels a $1.1M grant for his T32 Training proposal in Biomedical Big Data Sciences. The purpose of this pre-doctoral training program is for the trainee to become an expert in statistics or computer science (CS) or computational science, engineering, and mathematics (CSEM) or biology (via a PhD in neuroscience or ecology, evolution, and behavior [EEB] or cell and molecular biology [CMB]) while also having essential training in statistics and computer science and biology so they are best equipped to make important scientific discoveries using big data. The T32 training program breaks traditional graduate program boundaries by creating a formalized mechanism that allows students to cross disciplines during their PhD training by rotating among biomedical and computer science/statistics research groups. Such interdisciplinary training is essential for producing a biomedical workforce that is skilled in big data analysis. This is an exciting opportunity for the new PhD program in statistics and the already established PhD programs involved, and it is consistent with the interdisciplinary emphasis of the entire faculty involved with the proposal.

Faculty and Staff Awards
Dr. Kristin Harvey was recognized on April 5, 2016 by the College of Natural Sciences for her outstanding efforts and pursuit of excellence with the 2016 Teaching Excellence Award. Dr. Harvey teaches and coordinates SDS 302: Data Analysis for the Health Sciences and has been with the department as a lecturer since August of 2013.

Dr. James Scott was invited to give the Susie Bayarri Lecture during the 13th meeting of the International Society for Bayesian Analysis (ISBA) in Cagliari, Sardinia, Italy, during June 13–17, 2016. The Susie Bayarri Award is given every two years to recognize the research contributions of a Bayesian statistician under 35 years old, and the recipient is invited to give a plenary lecture at the world meeting of the International Society for Bayesian Analysis.

PhD in Statistics Students Enter Candidacy
During Fall 2016, three of the PhD in Statistics students passed their oral exam and formally entered candidacy. This is a milestone for the program as these are the first students in the program to reach this stage of their academic career. The candidates are:

• Guy Cole: supervised by Dr. Sinead Williamson
• Oscar Madrid Padilla: supervised by Dr. James Scott
• Tianjian Zhou: supervised by Dr. Peter Mueller

New Staff
The Department of Statistics and Data Sciences welcomed three new staff members to the administrative team this year, Rachel Poole (Sr. Administrative Associate) and Labrisha Betts (Grants and Development Coordinator) and Elif Ozdingis (Graduate Assistant, Spring 2016).

• Rachel Poole comes to SDS from UT’s central Human Resources office where she worked as an Administrative Associate. In her previous HR position, Rachel was awarded the President’s Exemplary Service Award in 2013.
• Labrishia Betts comes to UT from the Department of Aging and Disability Services and had previously held the position of Administrative Associate with UT’s Office of Sponsored Projects.
• **Elif Ozdingis** came to SDS from the Cockrell School of Engineering to assist in the planning, implementation, and logistical support of the 9th Annual UT Summer Statistics Institute and the NSF-CBMS Conference.

**Visiting Researchers**
The Department of Statistics and Data Sciences welcomed two visiting researchers this year, Lorenzo Capello and Daiane Zuanetti. Both were in residence in the department from September 2015 to May 2016.

• **Lorenzo Capello** comes from Bocconi University in Milan, Italy, where he is a Ph.D. student candidate. He received his bachelor's degree in Naval Engineering with a focus on applied mathematics and physics. He then continued his studies with a master's in Statistics and Finance, where he became passionate about time-series modeling. His areas of interest are Bayesian nonparametric inference, nonparametric regression, Markov models, and time-series analysis. During his residency at UT he worked with Professor Stephen Walker.

• **Daiane Zuanetti** comes from the Universidade Federal de São Carlos, Brazil, where she is a Ph.D. student. She earned both her bachelor's and master's degrees in Statistics from the Universidade Federal de São Carlos, and she subsequently began her Ph.D. in Statistics in March 2013 at that same institution after working for seven years in the financial market with credit and collection modeling in São Paulo, Brazil. Her areas of interest are Bayesian inference, computational methods (mainly MCMC) and statistical modeling for genetics data sets. During her residency at UT she worked with Professor Peter Mueller.

**SDS358: Applied Regression Analysis:**
This spring, Dr. Michael J. Mahometa offered a new course for undergraduate non-majors: SDS 358 Applied Regression Analysis. Designed as a “second course” in statistics, to be taken after the SDS302, SDS304/6, or SDS328M, the course was developed as a hybrid teaching platform, offering students two lectures per week to introduce a regression technique, with a single hands-on lab (lead by the instructor) to apply the technique using real world data.

To facilitate this structure, Dr. Mahometa developed each week’s lab and lab assessment with new data, along with custom scripts in R. To help student understanding of analytical techniques, Dr. Mahometa also created a custom R package for the course that included implementations of regression diagnostics, regression reporting, and computational tools such as the Johnson-Neyman technique for moderator significance. Students showed a culmination of learning through a final Poster Session. This Poster Project was of the student’s choosing, selecting their own research question, gathering the appropriate data, applying the correct regression technique, and then coherently offering an interpretation on the results to both the instructor and their peers. Dr. Mahometa plans to offer SDS358 Applied Regression Analysis again in Spring 2017.

**SDS 302/328M Final Poster Project**
Beginning in Fall 2015, two of the undergraduate statistics classes, SDS 302: Data Analysis for the Health Sciences and SDS 328M: Biostatistics, began a new format for their comprehensive, end-of-semester assessments. Rather than giving comprehensive final exams that covered content that had been tested on previously, these courses instituted a new final poster project. Between the Fall and Spring semesters (2015-2016) approximately 2500 students created research questions, collected data, and presented their poster projects during their final exam hours.
Students were given autonomy to select a topic of their choosing, but had to focus their studies on a research question that could be analyzed using specific techniques. Through these projects, students learned skills not taught in a traditional statistics lab that often includes data manipulated to meet statistical assumptions. Through this experience students have an opportunity to witness that, in reality, data is often messy and, even with a researcher’s best intentions and plans things occasionally do not work out as planned. This shows the true value of the project, allowing students the opportunity to reflect and learn from their mistakes to help them learn how to be better researchers in the future.

SDS Faculty Host first grant conference in Topological Data Analysis
Professors Lizhen Lin (SDS), Peter Mueller (SDS and Mathematics) and Rachel Ward (Mathematics) hosted the NSF-CBMS Regional Conference 2016: Topology, Geometry and Statistics at the UT Thompson Conference Center in June 2016. This conference introduced graduate students and junior researchers to topological data analysis (TDA), an active new field, which lies at the intersection of topology, geometry, and statistics. TDA offers the opportunity to foster research collaborations and has generated great interest across mathematics, statistics, computer science, machine learning, and electrical engineering communities. The conference was attended by 78 registered outside participants and 15 additional local participants (graduate students from statistics, mathematics and other departments).

Professor Sayan Mukherjee from Duke University was the principal lecturer. He delivered 10 lectures, including 1 data analysis session. In addition, the conference included five featured speakers: Professors Rabi Bhattacharya (University of Arizona), Susan Holmes (Stanford University), Ann Lee (Carnegie Melon University), Lek-Heng Lim (University of Chicago), and Yusu Wang (Ohio State University).

New Stat App Server:
In Spring 2016, SDS partnered with the College of Natural Sciences Information Technology Group, the Department of Mathematics, The University of Texas at Austin’s Academic Technology Support Group, and the College of Liberal Arts Instructional Technology Services, to bring a new Stat Apps Server online. Like the old server, access is available for all current University of Texas at Austin students, faculty, and staff.

The new server contains two Intel Xeon 2.5GHz 12-core processors, with 256GB of high speed RAM, capable of supporting hundreds of users simultaneously. Seats on the server can be used in support of academic instruction, for students to use software specific for class, or seats can be used by researchers to test-drive software such as Mplus or HLM7.

Resources Webpage
In response to the continued growth and increasing administrative demand on the department’s administrative staff, Administrative Manager Sasha Schellenberg created and implemented new administrative forms and a website resource page for SDS faculty, staff, and students. The new Resources section of the SDS website ensures that faculty, staff, and students are able to easily locate important information and documents. The page consists of three separate tabs: 1) an administrative resource page that contains access to policies and procedures for the department and the university, including copy and room requests; 2) an information technology resource page that contains important computing resources and contact numbers; and 3) a new faculty information resource page.
Disaster Recovery Plan
Under the directive of CNS Dean Linda Hicke, the department worked with the CNS Instruction Technology Office to develop a Disaster Recovery Plan. According to the Information Resources Use and Security Policy from The University of Texas at Austin, each college, school, or unit responsible for a computer or software system is directed to maintain a recovery plan that is reviewed periodically by the University’s Office of Internal Audits. The primary reason for a unit to engage in disaster recovery planning is to ensure the ability of the unit to function effectively in the event of a severe disruption to normal operations. Severe disruptions can arise from natural disasters, equipment failures, process failures, mistakes and errors in judgment or malicious acts. The department worked together to establish a disaster recovery plan that incorporates plans and procedures that will enable effective and timely recovery from the effects of disasters and events that result in data loss or data corruption, system failure and/or other operational failures.
II. UNIT REPORTS

A. INSTRUCTION

The demand for statistics and scientific computation courses continues to grow, particularly at the service course-level. SDS has responded by creating new courses and adding additional sections of in-demand courses to meet the needs of the diverse UT student population. SDS now offers more than 85 undergraduate and graduate courses each year, ranging from introductory data analysis and undergraduate honors statistics, to advanced graduate topics in Bayesian modeling and computational statistics.

Enrollment Trends

SDS faculty taught 63 undergraduate courses and 23 graduate courses to a total 5879 students in 2015–16—an 11.28% increase over last year.
Campus Usage
SDS courses serve a broad cross-section of students across campus. Enrollment in undergraduate SDS courses during 2015–16 is shown by college:

![Bar chart showing enrollment in undergraduate SDS courses by college]

Enrollment in graduate SDS courses during 2015–16 is shown by college:

![Bar chart showing enrollment in graduate SDS courses by college]

SOS Curriculum Innovation Grant
Following the success of the SDS 328M Course Transformation project in 2014, Dr. Kristin Harvey and Sally Amen applied for and had been awarded a second Curriculum Innovation Grant from the Center for Teaching and Learning. The grant funds the creation of the Statistics Online Support (SOS) resource platform. Designed for student researchers who have collected...
data but need guidance selecting and performing statistical analyses, SOS will provide on-demand content and tutorials that guide users to independently select the data analysis method appropriate for answering their research questions.

Dr. Harvey and Ms. Amen created a website to provide students a structured decision tree based platform where users will follow embedded links that will guide them to the types of statistical methods appropriate for their situation and tutorial videos on how to carry out those analyses. The site is undergoing final review before release for beta testing.

SOS will be a great resource for undergraduates in Freshman Research Initiatives (FRI) and Scientific Inquiry Across the Disciplines (SIAD) signature courses, as well as for graduate students working on quantitative research related to journal article submissions or their dissertations. In addition, students enrolled in introductory statistics courses will be able to use SOS as an additional study tool as much of the content on the site will be aligned with current undergraduate courses offered by SDS. Work on SOS will be completed during the Summer 2016, with an anticipated rollout during Fall 2016.

OnRamps
Since Fall 2013, Dr. Michael Mahometa and Sally Amen have worked alongside Senior Assistant Dean Cathy Stacy and with the Center for Teaching and Learning to transform SDS 302 Data Analysis for the Health Sciences into a course for The University of Texas OnRamps initiative. In its second full year of operation, the dual enrollment course partnered with 17 different Texas high schools, and 143 high school students earned college credit for SDS 302 on their transcript, with nearly half (48%) earning an “A.”

MOOC
In January of 2016, Dr. Michael J. Mahometa released a second run of the department’s Massively Open Online Course (MOOC), Foundations of Data Analysis (FoDA), through the edX platform. Unlike the first offering, this version provided online learners with a shortened two-part course: one for Descriptive Statistics and Visualizations, and another for Inferential Statistics. Both parts of FoDA were six weeks in length, with live office hours each week hosted through Google+ Hangouts OnAir. FoDA includes 12 total weeks of instruction, with over 40 instructional content videos, over 20 R tutorial videos, and a scaffolded learning structure that reliably guides students through the analysis process, while increasing independence. In addition to the past offering’s content, this current offering added an additional week of Inferential Statistics that focused on regression analysis. Students for the current release of FoDA also benefited from a new instructional tool: direct video feedback for incorrect items on all weekly PreLabs.

This second offering, like its predecessor, showed remarkable success. Over 31,000 students, representing 176 countries, enrolled in both parts of FoDA. The course also saw over 31% participant retention in Part One and over 50% participant retention in Part Two, based on video views in the last week of each part.

Instructional Videos Made Available
In Summer 2014, Dr. Kristin Harvey led the effort to share information and materials developed during past Course Transformation Projects to impact a broader audience beyond the students enrolled in the transformed courses. A new website now offers open-access to several video tutorials, providing a web resource that can be utilized by other undergraduate courses within the university and anyone outside the university. Previously, the course content designed for specific courses during the transformation was only accessible to students currently enrolled in
the courses. New videos are added to the website as new software and hand calculation instructional content has been created. The website can be viewed at stat.utexas.edu/videos.

With the addition of the Excel 2016 software a new set of videos were made and added to the site. These videos add to the already created and popular videos.

**Noteworthy:**
- Dr. Mahometa has been awarded the opportunity to release a third offering of Foundations of Data Analysis, again in the two-part format, in Fall 2016.
- To date, the instructional videos developed by Dr. Kristin Harvey and Sally Amen show the top video accumulating 543 views (“Top Ten Tips When Using Excel”), while 54 of the videos have had over 100 views to date.
B. CONSULTING

SDS provides free statistical consulting services to students, faculty, and staff. Clients receive assistance in planning and interpreting analyses, working with statistical software, planning research study designs, and learning how to better organize and manipulate their data. Faculty may also request contract consulting services for more in-depth analyses. In addition, SDS offers a collection of highly-subscribed instructional short courses in various programming languages and statistical software packages for a nominal fee.

Free Consulting

Free consulting services continued to be in high demand in 2015–16. SDS consultants provided 866 hours of free consulting to UT faculty, staff, and students, for an average of 17.3 hours per week. Clients were from almost all university colleges/schools, including the College of Natural Sciences, Moody College of Communication, College of Liberal Arts, College of Education, School of Architecture, and the College of Pharmacy. Student clients used 87% of these hours, while faculty and staff used the remaining 13%. Ninety-seven percent of consulting clients reported that their most recent consult was either “very good” or “excellent.” Nearly all 2015–16 consulting clients (99%) reported that they felt that their research results will be of higher quality because of the consulting service.

Contract Consulting

SDS provided 124 hours of contract consulting to UT faculty and organizational units and the Texas State Board of Podiatric Medical Examiners in 2015–16. Contract consulting clients generally have more complex data analysis needs and pay a fee for these services. SDS consultants typically provide a detailed reporting of results suitable for publication. Clients seeking contract consulting during the 2015–16 academic year included faculty and staff from across The University represented by the College of Liberal Arts, Department of Linguistics (27%), College of Education (16%), state agencies (8%), and the new Dell Medical School Program (8%).

In addition, SDS Contract Consulting worked with The University of Texas at Austin Graduate School to develop and implement the first University wide Graduate Student Housing Survey, which provided insight to incoming graduate students, when deciding on an area of Austin to live. Results are publically available on the Graduate School website.

Short Courses

In 2015–16, SDS taught a total of 29 short courses, covering various software packages such as R, Stata, Python, and Matlab, to 543 registrants. SDS offered four new short courses in Fall 2015, including an Excel software course as well as three courses in the new Topics in Statistics series taught by the statistical consulting staff. Instead of being focused on particular software, these topic courses cover a specific content area, ranging from introductory data analysis to advanced statistical methods.

Overall, SDS’s software and topic short courses saw high enrollment and positive student feedback in 2015–16. Ninety-eight percent of all short course participants said they would recommend the course they took to others.
**Noteworthy:**

- SDS Consulting services continued to be in high demand across the University, with over 99% of all 2015–16 available appointments filled and an average wait time in the Spring 2016 semester exceeding 10 days.
- Graduate student Cindy Blanco joined the consulting team in Spring and Summer 2016 as a student consultant. Cindy was a Graduate Fellow during Fall 2015.
C. SDS GRADUATE DEGREE PROGRAMS

Ph.D. in Statistics
The Ph.D. in Statistics program welcomed its third cohort of eight students in August 2015 and opened its application window for Fall 2016 in September 2015. The program received 138 completed applications for Fall 2016 admission. Fifteen offers of admission were made, resulting in an expected yield of five students.

Admissions Data: PhD Statistics

<table>
<thead>
<tr>
<th></th>
<th>Fall 2013</th>
<th>Fall 2014</th>
<th>Fall 2015</th>
<th>Fall 2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applied</td>
<td>79</td>
<td>104</td>
<td>130</td>
<td>138</td>
</tr>
<tr>
<td>Admitted</td>
<td>13</td>
<td>13</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Selectivity (%)</td>
<td>16.46</td>
<td>12.5</td>
<td>11.5</td>
<td>10.9</td>
</tr>
<tr>
<td>Enrolled</td>
<td>5</td>
<td>6</td>
<td>8</td>
<td>5*</td>
</tr>
<tr>
<td>Yield (%)</td>
<td>38.46</td>
<td>46.15</td>
<td>53.33</td>
<td>33.33</td>
</tr>
</tbody>
</table>

*Expected

PhD Student Progress

<table>
<thead>
<tr>
<th>Academic Year</th>
<th>Enrollment Count</th>
<th>Candidacy Count</th>
<th>Graduates</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015–2016*</td>
<td>16</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>2014–2015</td>
<td>11</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2013–2014</td>
<td>5</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Master's of Science in Statistics
The M.S. in Statistics program did not accept applications from terminal students this year as a faculty committee was tasked with reviewing and revising the program to address feedback provided by the external review committee.

The program continues to attract an increasing number of applicants for its concurrent cohort. Concurrent students are those who are admitted to the M.S. in Statistics program internally while pursuing a Ph.D. in another field such as Biological Sciences, Educational Psychology, Civil Engineering, and Sociology.

Enrollment Trends

<table>
<thead>
<tr>
<th>Academic Year</th>
<th>Enrollment Count</th>
<th>Graduates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M.S. Only</td>
<td>M.S./Ph.D.</td>
</tr>
<tr>
<td>2015–2016*</td>
<td>10</td>
<td>37</td>
</tr>
<tr>
<td>2014–2015</td>
<td>28</td>
<td>25</td>
</tr>
<tr>
<td>2013-2014</td>
<td>22</td>
<td>18</td>
</tr>
<tr>
<td>2012-2013</td>
<td>30</td>
<td>17</td>
</tr>
<tr>
<td>2011-2012</td>
<td>28</td>
<td>10</td>
</tr>
<tr>
<td>2010-2011</td>
<td>20</td>
<td>5</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>82</strong></td>
<td></td>
</tr>
</tbody>
</table>

*Enrollment and graduation numbers current through Spring 2016
Noteworthy:

- Four Professional Development (travel) Awards were provided to current PhD and MS in Statistics students to support a paper or poster presentation of their research. These awards were funded by the Graduate School or the department. The recipients of the awards were:
  - Carlos Pagani Zanini: to present a poster, “A Dynamic Partition Model for Reverse Phase Protein Arrays (RPPA) Data” at the Brazilian Meeting on Bayesian Statistics.
- Two PhD students have accepted publications:
D. GRADUATE FELLOWS PROGRAM

The Graduate Fellows program is a selective semester-long mentorship for UT graduate students that provides considerable training and experience in statistical analysis and consulting on applied problems in a variety of disciplines. Students learn new statistical methods and gain the confidence to teach themselves additional methods in the future. The students also build the interpersonal and presentation skills required of a professional consultant.

The Graduate Fellows take the SDS 388 Consulting Seminar taught by Dr. Michael Mahometa with the data analysis component advised by Dr. Matt Hersh. The course focuses on teaching the skills of statistical consulting, data analysis, and statistical methods. Additionally, the Graduate Fellows gain hands-on experience applying statistical methods using data from faculty and researchers across campus.

Fall 2015 Graduate Fellows
Selection for the program is a highly competitive process. For the Fall 2015 program, several graduate students from departments all over campus applied for five positions. To be eligible, students must have 30 hours of graduate study or a master’s degree and have a strong applied statistical background, including use of statistical software. The students selected were

- Cynthia Blanco (PhD student in Linguistics)
- Shihui Jia (PhD student in Operations Research & Industrial Engineering)
- Kejin Lee (PhD student in Educational Psychology)
- Nathaniel Pope (PhD student in Ecology, Evolution, & Behavior)
- Anastasiya Travina (MS student in Statistics)

Data Analysis Clients in Fall 2015
Data analysis is a fundamental component of the program. Fellows work with faculty members to provide assistance with the statistical or mathematical analysis of their research data. This allows students to gain hands-on experience applying statistical methods to real data while providing faculty and researchers across campus a valuable service. The Graduate Fellows worked with faculty members and NetSpend, one of our Corporate Partners.

Noteworthy:
- During SDS 388, each student was placed on a team that worked on a single project. This allowed students to dig deeper and get more involved with their projects in contrast to past years when students worked on multiple projects for the semester.
- At the end of the semester, the team working on NetSpend’s project gave a superb presentation of their results to an audience that included the president of the company. This was a tremendous learning experience for them.
- Cynthia Blanco, one of the Graduate Fellows, was selected to work as a consultant for the department’s consulting services during Spring and Summer 2016.
E. PORTFOLIO AND CERTIFICATE PROGRAMS

SDS provides four unique opportunities for students seeking to develop competencies in statistical modeling or scientific computation. The Portfolio in Applied Statistical Modeling and the Portfolio in Scientific Computation are 12-credit programs available to graduate students interested in strengthening and applying these skills to their research area. The Certificate in Scientific Computation and the Certificate in Applied Statistical Modeling are similar 18-hour programs available to undergraduate students.

Enrollment Trends

Portfolio in Applied Statistical Modeling
Since its inception in Fall 2009, 187 students have been admitted into the program and 86 students have completed the portfolio requirements.

<table>
<thead>
<tr>
<th>Academic Year</th>
<th>Enrollment Count</th>
<th>Completion Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015–2016*</td>
<td>40</td>
<td>17</td>
</tr>
<tr>
<td>2014–2015</td>
<td>39</td>
<td>15</td>
</tr>
<tr>
<td>2013–2014</td>
<td>47</td>
<td>16</td>
</tr>
<tr>
<td>2012–2013</td>
<td>54</td>
<td>15</td>
</tr>
<tr>
<td>2011–2012</td>
<td>43</td>
<td>11</td>
</tr>
<tr>
<td>2010–2011</td>
<td>24</td>
<td>11</td>
</tr>
<tr>
<td>2009–2010</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>86</strong></td>
<td></td>
</tr>
</tbody>
</table>

*Enrollment and graduation numbers current through Spring 2016

Participating students come from 10 colleges and schools across campus.

Student Participation Snapshot

<table>
<thead>
<tr>
<th>College/School</th>
<th># students currently enrolled</th>
<th># students completed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cockrell School of Engineering</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td>College of Communication</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>College of Education</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>College of Liberal Arts</td>
<td>11</td>
<td>21</td>
</tr>
<tr>
<td>College of Natural Sciences</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>Jackson School of Geosciences</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>LBJ School of Public Affairs</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>McCombs School of Business</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>School of Nursing</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>School of Social Work</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>40</strong></td>
<td><strong>86</strong></td>
</tr>
</tbody>
</table>
Portfolio in Scientific Computation
Since its inception in Fall 2010, 26 students have been admitted to the program and six students have completed the portfolio requirements.

<table>
<thead>
<tr>
<th>Academic Year</th>
<th>Enrollment Count</th>
<th>Completion Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015–16*</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>2014–2015</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>2013–2014</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>2012–2013</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>2011–2012</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>2010–2011</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>6</strong></td>
<td></td>
</tr>
</tbody>
</table>

*Enrollment and graduation numbers current through Spring 2016

Participating students come from three colleges and schools across campus.

Student Participation Snapshot

<table>
<thead>
<tr>
<th>College/School</th>
<th># students currently enrolled</th>
<th># students completed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cockrell School of Engineering</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>College of Education</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>College of Natural Sciences</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>5</strong></td>
<td><strong>6</strong></td>
</tr>
</tbody>
</table>

Certificate in Scientific Computation
Since its inception in Fall 2009, 200 students have been admitted into the program, and 32 students have completed the program.

<table>
<thead>
<tr>
<th>Academic Year</th>
<th>Enrollment Count</th>
<th>Completion Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015–2016*</td>
<td>69</td>
<td>6</td>
</tr>
<tr>
<td>2014–2015</td>
<td>69</td>
<td>4</td>
</tr>
<tr>
<td>2013–2014</td>
<td>79</td>
<td>5</td>
</tr>
<tr>
<td>2012–2013</td>
<td>71</td>
<td>6</td>
</tr>
<tr>
<td>2011–2012</td>
<td>51</td>
<td>6</td>
</tr>
<tr>
<td>2010–2011</td>
<td>27</td>
<td>4</td>
</tr>
<tr>
<td>2009–2010</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>32</strong></td>
<td></td>
</tr>
</tbody>
</table>

*Enrollment and graduation numbers current through Spring 2016

Participating students come from four colleges and schools across campus.

Student Participation Snapshot

<table>
<thead>
<tr>
<th>College/School</th>
<th># students currently enrolled</th>
<th># students completed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cockrell School of Engineering</td>
<td>18</td>
<td>10</td>
</tr>
<tr>
<td>College of Liberal Arts</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>College of Natural Sciences</td>
<td>41</td>
<td>18</td>
</tr>
<tr>
<td>Jackson School of Geosciences</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>McCombs School of Business</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>69</strong></td>
<td><strong>32</strong></td>
</tr>
</tbody>
</table>
Certificate in Applied Statistical Modeling
Since its inception in Fall 2013, 67 students have enrolled and 11 students have completed the program.

<table>
<thead>
<tr>
<th>Academic Year</th>
<th>Enrollment Count</th>
<th>Completion Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015–2016*</td>
<td>56</td>
<td>9</td>
</tr>
<tr>
<td>2014–2015</td>
<td>27</td>
<td>1</td>
</tr>
<tr>
<td>20132014</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td>11</td>
</tr>
</tbody>
</table>

*Enrollment and graduation numbers current through Spring 2016

Participating students come from six colleges and schools across campus.

Student Participation Snapshot

<table>
<thead>
<tr>
<th>College/School</th>
<th># students currently enrolled</th>
<th># students completed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cockrell School of Engineering</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td>College of Communication</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>College of Liberal Arts</td>
<td>16</td>
<td>3</td>
</tr>
<tr>
<td>College of Natural Sciences</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>McCombs School of Business</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>Undergraduate Studies</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>TOTAL</td>
<td>56</td>
<td>11</td>
</tr>
</tbody>
</table>

Noteworthy:
- Dr. Robin Gutell, Program Director for the Certificate in Scientific Computation and Dr. Mia Markey, the research advisor updated the research requirements and process for the program. The new guidelines and examples should assist students in completing the research component of the program.
- Starting Fall 2016, the name of the Certificate in Scientific Computation will become the Certificate in Scientific Computation & Data Sciences.
- Graduates of the certificate, portfolio, and MS in Statistics programs were recognized during the 4th annual SDS Spring Student Celebration.
F. SDS SEMINAR SERIES & STATISTICS IN ACTION

SDS Seminar Series
Since Fall 2011, the SDS Seminar Series has made a vital contribution to the intellectual, cultural, and scholarly environment at The University of Texas at Austin for students, faculty, and the wider community. The lecture series provides participants with the opportunity to hear from leading scholars and experts who work in different applied areas, including business, biology, medicine, computer modeling, and economics. Each talk is free of charge and open to the public. The Fall 2015 SDS Seminar Series featured six speakers with an average attendance of 34 people. The Spring 2016 SDS Seminar Series featured 14 speakers with an average attendance of 40. (Please see Appendix C for program details.)

Statistics in Action
Since Fall 2009, the Statistics in Action series has drawn the interest of undergraduate students enrolled in the introductory statistics courses offered by SDS. The series provides an opportunity to expose undergraduate students to how professionals in various fields use data. During the Fall 2015 and Spring 2016 semesters, the Statistics in Action series included seven participating organizations:

- Freescale Semiconductor (now NXP)
- Allergan Pharmaceuticals
- Mass Relevance
- Office of the State Demographer
- Unique Influence
- Netspend
- Dell Computers, Inc.

Overall, student participation was consistently high throughout the 2015–2016 school year. In Fall 2015, the overall attendance was 1036 students with a weekly average of about 173 students across six events. The Spring 2016 series brought in 1,023 students with a weekly average of approximately 146 students across seven events. Instructors of introductory courses continue to offer incentives for students attending these talks, which keeps attendance consistent. Several of the instructors require students to submit a one-page write-up explaining what they learned from the talk. This has helped students make connections between the talks and their own learning. After several of the talks students spoke with the presenters about opportunities for internships, furthering the extracurricular benefit of the sessions to the students. Students are primarily from the College of Natural Sciences, the College of Communication, the School of Nursing, and the College of Liberal Arts with a total of eleven colleges and schools across the university represented. (Please see Appendix D for program details.)
G. SUMMER STATISTICS INSTITUTE

The 9th annual UT Summer Statistics Institute (SSI) was held May 23–26, 2016, in collaboration with Academic Technology Support in the College of Liberal Arts building. SSI provides a unique hands-on opportunity for participants to acquire valuable skills directly from experts in the field. Participants joined the Institute from across the country, coming from as far away as Washington and Massachusetts, illustrating the growing popularity of the UT Summer Statistics Institute nationwide. SSI featured 24 twelve-hour courses designed to appeal to a broad range of students, faculty, staff, and the public. New courses this year included “Introduction to Data Science in Industry with R” (Richard Leu), “Introduction to Data Science in Python” (Tal Yarkoni), “Introduction to SQL and Relational Database Design” (Chris Golubski), and “Non-Parametric Statistical Methods for Small Data Sets” (Bindu Viswanathan).

The 2016 SSI saw a record student enrollment of 689 participants (a 13% increase over 2015’s enrollment), with 24 participating instructors from 15 departments. This year’s attendance breakdown was: 33% UT students, 22% UT faculty and staff, 8% non-UT students, and 37% non-UT other (e.g., private industry, state departments, and non-profits).

The 2016 SSI brought in a record gross income of $211,685.00 with a projected total net profit of $113,770.00, a 20% increase over prior year.

Noteworthy:
- 45% of SSI enrollees (303 of the 689 participants) came from outside The University of Texas at Austin. This 8% increase in outside enrollment over the 2015 SSI is likely due to increased awareness of SSI through targeted advertising and continued participant satisfaction (word of mouth).
- 87% of SSI participants who completed the course evaluations said that they would recommend SSI to others.
- 89% of SSI participants who completed the course evaluations said that the instructor increased their knowledge and competence in the area of study.
H. CORPORATE PARTNERSHIPS

SDS initiated the Corporate Partnership Program in Spring 2013. Through this program, local corporations work together with the department to enrich the educational experience of our students. As part of the program, the corporate sponsors provide actual datasets and problems for the students.

Students gain practical experience while working with professionals to apply statistical methods they have learned in class to real-world data. Students are involved in all aspects of the data analysis to give them a real-world consulting experience: from discussing the goals and details of the experiments, to analyzing the data, and preparing and presenting the results.

This past Fall, we had a very successful collaboration with NetSpend. The data analysis was run in conjunction with the consulting seminar and the Graduate Fellows program. A team of five students was assigned to work on the project but many more students provided input and support. The project culminated in a presentation in the offices of NetSpend by eight students from the consulting seminar.

Noteworthy:
- The students made great strides in resolving a real-word problem facing a major local corporation and did a tremendous job presenting their results.
I. GRANTS

During 2015–16, SDS assisted faculty with the submission of more than eight individual faculty-driven research proposals to such agencies as:

- National Institutes of Health (NIH)
- National Science Foundation (NSF)
- Naval Postgraduate School

_and in collaboration with_

- NorthShore University Health System
- University of Pennsylvania
- University of Maryland
- Harvard University
- University of Florida
- Brown University
- Boston University
- Carnegie Mellon University
- University of Tennessee Health Science Center
- University of Minnesota

SDS had 14 active grants and contracts this fiscal year:

<table>
<thead>
<tr>
<th>Faculty</th>
<th>Active Grants &amp; Contracts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michael Daniels, Professor</td>
<td>• PI: Bayesian Approaches for Missingness and Causality in Cancer and Behavior Studies, NIH R01</td>
</tr>
<tr>
<td></td>
<td>• PI: Predoctoral Training in Biomedical Big Data Science, NIH</td>
</tr>
<tr>
<td></td>
<td>• Sub-award PI: Magnetic Resonance Imaging and Biomarkers for Muscular Dystrophy, NIH</td>
</tr>
<tr>
<td></td>
<td>• Sub-award PI: Hospital Responses to Medicare Readmission Penalties, NIH</td>
</tr>
<tr>
<td></td>
<td>• Sub-award PI: Rural Lifestyle Eating and Activity Program (Rural LEAP), NIH/National Heart, Lung, and Blood Institute</td>
</tr>
<tr>
<td></td>
<td>• Sub-award PI: Non-Parametric Bayesian Methods for Causal Inference, NIH/National Institute of General Medical Sciences</td>
</tr>
<tr>
<td></td>
<td>• Sub-award PI: Optimizing HIV Treatment Monitoring under Resource Constraints, NIH</td>
</tr>
<tr>
<td>Lizhen Lin, Assistant Professor</td>
<td>• PI: CBMS Conference: Topological Data Analysis: Topology, Geometry and Statistics, NSF</td>
</tr>
<tr>
<td></td>
<td>• Co-PI: BIGDATA:F:DKA: Statistical Foundations for Analyzing Large Collections of Network Data Objects, ARO</td>
</tr>
<tr>
<td></td>
<td>• Subaward PI: Mathematical Foundations for Analyzing Large</td>
</tr>
<tr>
<td>Professor/Professor</td>
<td>Collections of Combinatorial-Data Objects, Boston University</td>
</tr>
<tr>
<td>---------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>Lauren Ancel Meyers, Professor</td>
<td>• Co-PI: Surety BioEvent App: BioEvent Surveillance, Detection and Prediction Leveraging Trusted NextGen Data Sources, Department of Defense (DOD)/Defense Threat Reduction Agency (DTRA)</td>
</tr>
</tbody>
</table>
| Peter Müller, Professor | • Co-Investigator: *Bayesian Approaches for Missingness and Causality in Cancer and Behavior Studies*, NIH R01  
• Sub-award PI: *Bayesian Inference for Tumor Heterogeneity with Next-Generation Sequencing Data*, NIH R01 |
| James Scott, Assistant Professor | • PI: *CAREER: Bringing Richly Structured Bayesian Models into the Discrete-data Realm via New Data-augmentation Theory and Algorithms*, NSF  
• Co-Investigator: *Bayesian Approaches for Missingness and Causality in Cancer and Behavior Studies*, NIH R01 |
| Sinead Williamson, Assistant Professor | • Co-PI: *BIGDATA:F:DKA: Collaborative Research: Theory and Algorithms for Parallel Probabilistic Inference with Big Data, via Big Model, in Realistic Distributed Computing*, NSF |

**Noteworthy:**
- Professor and Chair, Michael Daniels received funding from NIH for his T-32 Training grant proposal: Predoctoral Training in Biomedical Big Data Science. This training initiative is a collaboration with other departments on campus including, Computer Science, Biology, and Biomedical Engineering. Trainees in this program will become experts in statistics, computer science, or biology while also having essential training in all areas so they are best equipped to make important scientific discoveries involving big data.
J. DEVELOPMENT

SDS has partnered with the College of Natural Sciences (CNS) External Relations team to create both short- and long-term plans for building philanthropic and community partnerships with alumni, friends, corporations, and foundations. The second annual SDS newsletter was distributed to over 2000 recipients in November 2015, and SDS Development personnel met with industry representatives from companies such as Dell, PayPal, Cloudera, and Continuum Analytics to discuss partnering opportunities. In 2015, SDS received support from StataCorp LP for the department’s annual UT Summer Statistics Institute (SSI) and also received support from The National Science Foundation (NSF) for a conference on Topological Data Analysis that was hosted in June 2016. SDS is continuing to actively pursue opportunities for support of programs and conferences through NIH, NSF, and the Institute of Mathematical Statistics Meeting sponsorship.
Appendix A: SDS Organizational Chart

Dean
College of Natural Sciences
Linda Hicke, Ph.D.

Chair
Mike Daniels, Sc.D.

Assistant Director for Administration
Vicki L. Keller, MA
Coordinates daily SDS operations & student programs
Supervise departmental staff
Graduate Coordinator
Budgets & EOM
Communications
Reporting
FAR

Consulting Manager
Michael J. Mahometa, Ph.D.
Manages all consulting operations
Oversees FAC 101B & Stat App Server
On-ramps & MOOC development

Specialist
Kristin Harvey, Ph.D.
Coordinates SDS 302 & Statistics in Action
Teaches undergraduate courses

Lecturers
Lauren Blondeau, PhD
Sarah Collins, PhD
Steven Hernandez, MS
Matt Hersh, PhD
Jerry Manheimer, PhD
Maggie Myers, PhD
Mary Parker, PhD
Lindsey Smith, PhD
Bindu Viswanathan, PhD

SDS Core Faculty
(50% Time)
Michael Daniels, Sc.D. Integrative Biology
Peter Müller, Ph.D. Mathematics
Pradeep Ravikumar, Ph.D. Computer Science
Stephen Walker, Ph.D. Mathematics
Carlos Carvalho, Ph.D. IROM
James Scott, Ph.D. IROM
Sinead Williamson, Ph.D. IROM

SDS Full Faculty
(100% Time)
Lizhen Lin, Ph.D.
Purnamitra Sarkar, Ph.D.

Administrative Manager
Sasha Schellenberg, B.A.
Supervise classified staff
Course scheduling & catalog management
CIS, ALEKS, FASET, Final Exam scheduling, HB 2504 admin
Coordinates SDS Seminar Series & Summer Statistics Institute
Event planning

SDS Core Faculty
(0% Courtesy Appt)
Paul Damien, Ph.D. IROM
Lauren Meyers, Ph.D. Integrative Biology
Tom Sager, Ph.D. IROM
Tom Shively, Ph.D. IROM
Mingyuan Zhou, Ph.D. IROM

Student Assistant
Alexandra Fisk

Administrative Associate
Labrisha Betts
Grants: pre- and post-award processing Development

Senior Administrative Associate
Rachel Poole
HR processing
Purchasing
Payroll
Accounting
Inventory
Records Maintenance
Billing
DEFINE Vouchers

Staff Consultant
Sally Amen, M.S.
Coordinates SDS 328 Manages short courses On-ramps development

Staff Consultant (p/t)
Erika Hale, M.S.

Department of Statistics & Data Sciences 23
Appendix B: SDS Core Faculty

Carlos Carvalho, Associate Professor, IROM

Dr. Carvalho’s interest is in the development of methodological aspects of structured probability models for large-scale multivariate problems, with applications ranging from financial time series to high-throughput cancer genomics. His work pays special attention to the development and improvement of associated computational tools for model selection and inference, with current projects in financial econometrics and empirical asset pricing problems.

Dr. Carvalho did his undergraduate studies in Brazil. He received his Ph.D. in Statistics from Duke University in 2006. Before moving to The University of Texas, he was an Assistant Professor at The University of Chicago Booth School of Business.

Honors:
- Donald D. Harrington Faculty Fellow – The University of Texas, Austin, 2009–10
- IBM Corporation Scholar – The University of Chicago, 2008–09
- Leonard J. Savage Award for outstanding doctoral dissertation in Bayesian econometrics and statistics – Honorable Mention, 2006

Michael Daniels, Professor, Integrative Biology

Dr. Daniels’ research program revolves Bayesian methods for biostatistics with special attention to incomplete data, estimation of the dependence structures, and methods for causal inference. His current collaborations include clinical trials in weight management (which motivates development of causal methods for mediation) and muscular dystrophy (which motivates development of complex latent variable methods), and questions involving the impact of recent Medicare legislation on ‘preventable’ hospital outcomes.

Dr. Daniels did his undergraduate studies at Brown University and received his doctoral training in biostatistics at Harvard University in the early 1990’s. Most recently, he was Professor and Chair in the Department of Statistics at the University of Florida. Before that, Dr. Daniels spent five years on the faculty at Iowa State University and two years at Carnegie Mellon University.

Honors:
- Editor of Biometrics, 2015–2017
- The Lagakos Distinguished Alumni Award, Department of Biostatistics, Harvard University, 2014
- Fellow of the American Statistical Association, 2007
Lizhen Lin, Assistant Professor

Dr. Lin’s general areas of research include Bayesian nonparametric theory and asymptotics, statistical inference on manifolds and general spaces, shape constrained inference, big data analysis, and large sample theory for networks. Some of her ongoing theoretical work focuses on (1) building theory and methodologies for inference on manifold valued data such as density estimation and regression on manifolds; (2) developing both Bayesian and frequentist methods for shape constrained estimation and hypothesis testing; (3) developing scalable and robust procedures for Bayesian inference; and (4) establishing a large sample theory (such as deriving central limit theorems) for network valued data.

Dr. Lin received a Ph.D in Mathematics in 2012 from the University of Arizona under the guidance of Rabi Bhattacharya. Prior to that, she studied Mathematics and Statistics at Sichuan University from 2002–2006. Dr. Lin spent two years as a postdoctoral scholar from 2012 to 2014 at Duke University, mainly working with David Dunson. She has also been a member of the Laboratory for Psychiatric Neuroengineering in Duke University medical center since early 2013 working on neuro-psychiatric research.

Honors:
- Big Data Information Initiative at Duke (iiD) Research Incubator Award Grant, 2013
- Galileo Circle Scholar, 2011

Peter Müller, Professor, Mathematics

Dr. Müller’s interest is in methods and applications of Bayesian inference. More specifically, he is working on nonparametric Bayesian inference, decision problems, and applications to biomedical research problems. Nonparametric Bayesian inference refers to prior models for infinite dimensional random quantities, typically random probability measures. Decision problems include particular clinical trial design and multiple comparison procedures. Other applications that interest Dr. Müller include inference related to dependence structure, specifically graphical models to formalize inference about dependence for high throughput genomic data. Another large area of application is population pharmacokinetic and pharmacodynamic models, which give rise to many good applications that exploit many of his methodological interests.

Dr. Müller’s undergraduate education is from Universität Wien and Technische Universität Wien, Austria. He received his Ph.D. from Purdue University where he worked under Jim Berger on MCMC for constrained parameter problems. He spent several years at the Institute of Statistics and Decision Sciences (ISDS), Duke University, and at M.D. Anderson Biostatistics.

Honors:
- Fellow of the American Statistical Association
- President of the International Society for Bayesian Analysis, 2010
Pradeep Ravikumar, Associate Professor, Computer Science

Dr. Ravikumar's main area of research is in statistical machine learning. The core problem here combines the statistical imperative of inferring reliable conclusions from limited observations or data with the computational imperative of doing so with limited computation. Of particular interest are modern settings where the dimensionality of data is high, and simultaneously achieving these twin objectives is difficult. His recent research has been on the foundations of such statistical machine learning, with particular emphasis on graphical models, high-dimensional statistical inference, and optimization.

Dr. Ravikumar received his BTech in Computer Science and Engineering from the Indian Institute of Technology, Bombay in India. He received his Ph.D. in Machine Learning from the School of Computer Science at Carnegie Mellon University, where he worked with John Lafferty. He was a postdoctoral scholar at the Department of Statistics, University of California, Berkeley from 2007–2009, where he worked with Martin Wainwright and Bin Yu.

Honors:
- NSF CAREER Grant, 2012
- Honorable Mention, ACM SIGKDD Dissertation Award, 2008
- Honorable Mention, CMU School of Computer Science Distinguished Dissertation Award, 2007/08
- Siebel Scholar, 2007
- Indian National Talent Search Scholar

Purnamrita Sarkar, Assistant Professor

Dr. Sarkar works on large scale statistical machine learning problems with a focus on statistical models, asymptotic theory and scalable inference algorithms for large networks.

Dr. Sarkar graduated from the School of Computer Science at Carnegie Mellon University in 2010. After earning her doctorate she was a postdoctoral scholar at U. C. Berkeley jointly in the Department of Electrical Engineering and Computer Sciences and the Department of Statistics. She received her Bachelor's degree in Computer Science from the Indian Institute of Technology, Kharagpur in 2004.

Honors:
- Best paper award, 29th International Conference on Data Engineering (ICDE), 2013
- Best paper award, International Conference on Social Networks Analysis and Mining (ASONAM), 2009
James G. Scott, Associate Professor, IROM

Dr. Scott's research focuses on statistical methodology for high-dimensional data sets, with applications in a diverse set of areas spanning the social, physical, and biomedical sciences. Three areas of methodological focus include (1) large-scale multiple testing, anomaly-detection and screening problems, where the rate of false discoveries must be controlled in order to yield viable inferences; (2) inference in sparse models; and (3) the application of data-augmentation theory and algorithms to improve the efficiency of Bayesian inference in large-scale models for discrete data sets. His recent applied work has included collaborations in health care, demography, linguistics, biology, and neuroscience.

Dr. Scott received his Ph.D. from Duke University, where he studied Bayesian model selection under Jim Berger. Before that he studied at Trinity College, Cambridge for two years. He was an undergraduate from 2000 to 2004 at UT-Austin in the Dean's Scholars and Plan II honors programs.

Honors:
- Regents’ Outstanding Teaching Award, 2014
- NSF CAREER Grant, 2013
- Savage Award, 2010 (awarded by the International Society of Bayesian Statistics for best thesis in Bayesian statistical theory)
- National Science Foundation Graduate Research Fellowship, 2006–2009

Stephen G. Walker, Professor, Mathematics

Dr. Walker’s main research focus is on Bayesian parametric and nonparametric methods. He has worked on applications, methodology, theory, implementation via MCMC, and foundational issues. Dr. Walker’s main areas of applications include medical statistics and financial data. Recent work on Bayesian nonparametrics includes constructing time series and regression models. Recent work also includes working with Bayesian models under misspecification and using loss functions as an alternative to probability models within a learning process akin to Bayesian updating.

Dr. Walker received his BA (Hons.) in Mathematics at the Oriel College of Oxford University, being awarded Open Exhibition on entry to the college. He received his Ph.D. in Statistics from the Imperial College of London in 1995, supervised by Jon Wakefield. Dr. Walker has taught at various institutions: Imperial College at London, the University of Bath, and most recently at the University of Kent.

Honors:
- Chair of Bayesian Nonparametric Section of ISBA, 2010–2012
- EPSRC Advances Research Fellow, 2001–2006
Dr. Williamson’s main research focus is the development of nonparametric Bayesian methods for machine learning applications. In particular, she is interested in constructing distributions over correlated measures and structures, in order to model correlated data sets or data with spatio-temporal dependence. Examples include models for documents whose topical composition varies through time, and models for temporally evolving social networks. A key research goal is the development of efficient inference algorithms for such models, and she is currently investigating methods that allow us to apply Bayesian nonparametric techniques to large datasets.

Dr. Williamson received her MEng from the University of Oxford, MSc from University College London, and PhD from the University of Cambridge. Before joining the faculty at UT Austin, Sinead was a postdoctoral scholar at Carnegie Mellon University.
### Appendix C: SDS Seminar Series Speakers

<table>
<thead>
<tr>
<th>Date</th>
<th>Speaker Name</th>
<th>Institution</th>
<th>Title of Talk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sep. 11,  2015</td>
<td>Giovanni Petris</td>
<td>University of Arkansas</td>
<td>“A Framework for Functional Time Series Analysis”</td>
</tr>
<tr>
<td>Sep. 25,  2015</td>
<td>Alexander Volfovsky</td>
<td>Harvard University</td>
<td>“Testing and Estimation for Relational Data”</td>
</tr>
<tr>
<td>Oct. 16,  2015</td>
<td>Babak Shahbaba</td>
<td>University of California, Irvine</td>
<td>“Scalable Monte Carlo Methods”</td>
</tr>
<tr>
<td>Nov. 6,   2015</td>
<td>Fernando Perez-Cruz</td>
<td>Alcatel-Lucent, Bell Labs</td>
<td>“Infinite Factorial Dynamical Models/Bayesian Nonparametrics for Marathon Modeling”</td>
</tr>
<tr>
<td>Nov. 13,  2015</td>
<td>Larry Carin</td>
<td>Duke University</td>
<td>“Deep Poisson Factor Modeling”</td>
</tr>
<tr>
<td>Jan. 26,  2016</td>
<td>Veronika Rockova</td>
<td>Wharton School of the University of Pennsylvania</td>
<td>“Fast Bayesian Factor Analysis via Automatic Rotations to Sparsity”</td>
</tr>
<tr>
<td>Feb. 05,  2016</td>
<td>Amanda Mejia</td>
<td>Johns Hopkins University</td>
<td>“Statistical Methods for Brain Connectivity Analysis”</td>
</tr>
<tr>
<td>Feb. 12,  2016</td>
<td>Yen-Chi Chen</td>
<td>Carnegie Mellon University</td>
<td>“Statistical Inference using Geometric Features”</td>
</tr>
<tr>
<td>Feb. 19,  2016</td>
<td>Scott Berry</td>
<td>Berry Consultants</td>
<td>“Bayesian Advances in Clinical Trials: Platform Trials, Basket Trials, and Disease Progression Modeling”</td>
</tr>
<tr>
<td>Mar. 25,  2016</td>
<td>Karen Kafadar</td>
<td>University of Virginia</td>
<td>“Distinguishing ‘Typical’ from ‘Exotic’ in Streaming Data Sets”</td>
</tr>
<tr>
<td>Apr. 01,  2016</td>
<td>Stanislav Minsker</td>
<td>University of Southern California</td>
<td>“Efficient Representation of Data on Smooth Manifolds: Non-asymptotic Bounds and Robustness”</td>
</tr>
<tr>
<td>Apr. 08,  2016</td>
<td>Mike West</td>
<td>Duke University</td>
<td>“Structured Dynamic Graphical Models &amp; Scaling Multivariate Time Series Methodology”</td>
</tr>
<tr>
<td>Apr. 15,  2016</td>
<td>David Firth</td>
<td>University of Warwick</td>
<td>“How To Get Along With Sum 1: Generalized Linear Models for Compositional Data”</td>
</tr>
<tr>
<td>Apr. 29,  2016</td>
<td>Subhashis Ghoshal</td>
<td>North Carolina State University</td>
<td>“Bayesian Methods for Boundary Detection in Images”</td>
</tr>
<tr>
<td>Jun. 02,  2016</td>
<td>Swaminathan Vishwanathan</td>
<td>University of California, Santa Cruz</td>
<td>“Scaling up Bayesian Models”</td>
</tr>
</tbody>
</table>
### Appendix D: Statistics in Action Attendance, 2015–2016

<table>
<thead>
<tr>
<th>Semester</th>
<th>Attendance</th>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall 2015</td>
<td>193</td>
<td>Allergan Pharmaceuticals</td>
</tr>
<tr>
<td>Fall 2015</td>
<td>174</td>
<td>Unique Influence</td>
</tr>
<tr>
<td>Fall 2015</td>
<td>172</td>
<td>NetSpend</td>
</tr>
<tr>
<td>Fall 2015</td>
<td>163</td>
<td>Freescale</td>
</tr>
<tr>
<td>Fall 2015</td>
<td>168</td>
<td>Office of the State Demographer</td>
</tr>
<tr>
<td>Fall 2015</td>
<td>166</td>
<td>Spredfast</td>
</tr>
<tr>
<td>Spring 2016</td>
<td>152</td>
<td>Dell</td>
</tr>
<tr>
<td>Spring 2016</td>
<td>177</td>
<td>Allergan Pharmaceuticals</td>
</tr>
<tr>
<td>Spring 2016</td>
<td>140</td>
<td>Spredfast</td>
</tr>
<tr>
<td>Spring 2016</td>
<td>138</td>
<td>NetSpend</td>
</tr>
<tr>
<td>Spring 2016</td>
<td>154</td>
<td>Office of the State Demographer</td>
</tr>
<tr>
<td>Spring 2016</td>
<td>109</td>
<td>NXP</td>
</tr>
<tr>
<td>Spring 2016</td>
<td>153</td>
<td>Unique Influence</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>2059</strong></td>
<td></td>
</tr>
</tbody>
</table>