Curriculum Vitae of

# GUTTI JOGESH BABU

## Director, Center for Astrostatistics

http://astrostatistics.psu.edu


**Address:** 417C Joab L. Thomas Building,
Department of Statistics,
The Pennsylvania State University,
University Park, PA 16802.

**E-mail:** babu@psu.edu
**Telephone:** (814) 863-2837 **Fax:** (814) 863-7114
**Citizenship:** United States of America


**Employment:**

| | |
|---|---|
| 1985–present | Professor, Pennsylvania State University. |
| January 1982–85 | Professor, Indian Statistical Institute, India. |
| 1976–December 1981 | Associate Professor, Indian Statistical Institute, India. |

**Education:**

| | |
|---|---|
| 1970 | M. Stat., Indian Statistical Institute, Calcutta, India. |
| 1974 | Ph. D., Indian Statistical Institute, Calcutta, India. |

**Honors:**

Elected Fellow - Institute of Mathematical Statistics.
Elected Fellow - American Statistical Association.
Elected Fellow - American Association for the Advancement of Science.
Elected Member - International Statistical Institute.
National Research Council's Twinning Fellowship for 1997-1999.
Research Professor, Mathematical Sciences Research Institute, Berkeley (Jan.–Mar. 2005).
Mid-Atlantic Region University Continuing Education Association award for exemplary non-credit program development (Summer school in statistics for astronomers), 2007.

**Visiting Appointments:**

| | |
|---|---|
| Jan.-May. 2006 | SAMSI University Fellow, Statistical and Applied Mathematical Sciences Institute, Research Triangle Park, NC, and University of North Carolina, Chapel Hill |
| Jan. 2005 | University of Hyderabad, India |
| Dec. 1998 | Indian Statistical Institute, India |
| Oct.–Nov. 1998 | Rutgers University |
| Sep.–Oct. 1998 | Concordia University, Canada |
| June 1995 | National Sun Yet-sen University, Kaohsiung, Taiwan |
| May 1995 | Indian Statistical Institute |
| March 1992 | Mathematical Sciences Research Institute, Berkeley |
| Nov. 1991 | Monash University, Australia |
| Fall 1982 | University of Ottawa |
| 1981–1982 | Rutgers University |
| Spring 1981 | University of Arizona |
| Nov.–Dec. 1980 | Math. Inst. of the Hungarian Academy of Sciences, Budapest |

| | |
|---|---|
| Summer 1980 | University of Ottawa |
| 1975–1976 | University of Oregon |
| 1974–1975 | University of Illinois at Urbana-Champaign |
| 1973–1974 | Tata Institute of Fundamental Research, India |

**Editorial and Professional Service:**

Editor-in-Chief, *Statistical Methodology* (2003 - ).
Coordinating Editor, *Journal of Statistical Planning and Inference* 1998 - 2004, and Associate Editor 1995 - 1997.
Associate Editor, *Journal of Nonparametric Statistics* 1989 - 2007.
Co-Editor, *Sankhyā*, Series A 1981 - 1993 and 1999 - 2007.
Advisory Committee for *Sankhyā* 1993 - 1999.
Search committee to select editors for the journal Sankhya 1998.

*Institute of Mathematical Statistics* (**IMS**):
Chair, IMS Committee on **Fellows** 2002 - 2003.
Member, IMS Committee on Fellows 2001 - 2004.

*American Statistical Association* (**ASA**):
Member, ASA Committee on Scientific Freedom and Human Rights (January 1, 2010 to December 31, 2012).

*International Statistical Institute* (**ISI**):
Vice-Chair, Executive Board, ISI Astrostatistics Committee.
`http://isi-web.org/com/ast`

*Statistical and Applied Mathematical Sciences Institute* (**SAMSI**):
Chair, Program Leaders Committee, SAMSI Astrostatistics Program, January - May 2006.
Will be leading the second SAMSI program on Astrostatistics in 2012.

*Large Synoptic Survey Telescope* (**LSST**), `http://lsst.org`:
Core Team Member, **LSST** *Informatics and Statistics Science Collaboration* (2009 - ).
Member, **LSST** *Weak Lensing Science Collaboration team* (2006 - ).

*Summer Schools in Statistics for Astronomers*:
Organized annual *Summer Schools in Statistics for Astronomers* since 2005 at Penn State. Also organized these summer schools in collaboration with the *Indian Institute of Astrophysics* in July 2007, July 2008, and July 2010, at Vainu Bappu Observatory located near the village of Kavalur in India.

*Astrostatistics School*:
Organized an astrostatistics school for the Instruments Division at the Space Telescope Science Institute (STScI, http://www.stsci.edu/portal/) in September - November 2011.

*International Indian Statistical Association* (IISA): Program Chair 1998 - 2000.
Editorial Board, IISA Conference Proceedings Volume.

Served on NSF and NASA panels, including:
NSF Cyber-Enabled Discovery and Innovation Panel, NSF Software Infrastructure Panel, NASA Postdoc panel, NASA AISR Panel, NSF Mathematical Innovations in Astronomy Panel, NSF Mathematical Sciences Postdoctoral Research Fellowships panel, Member of Executive committee for NSF MSP Postdoc panel.

Reviewed research proposals for *NSF*, *NSA*, *Air Force Office of Scientific Research* and *Natural Sciences and Engineering Research Council of Canada*.

Written over 215 reviews for *Mathematical Reviews* and *Zentralblatt*.

Referee for the journals: Annals of Probability, Annals of Statistics, Sankhyā, Probability Theory and Related Fields, Journal of the American Statistical Association, Journal of Multivariate Analysis, Journal of Environmental and Ecological Statistics, Annals of the Institute of Statistical Mathematics, Communications in Statistics, Canadian Journal of Statistics, Journal of Statistical Planning and Inference, Econometric Theory, Indian Journal of Statistics, The Astrophysical Journal, Statistics and Probability Letters, Sociological methodology, Technometrics.

**Research Interests:**

Bootstrap and other resampling methods. Statistical applications to Astronomy and Physics. Analysis of massive data. Nonparametric Methods. Inference for misspecified models. Goodness-of-fit tests when parameters are estimated. Edgeworth expansions. Statistical Group Theory and its applications. Inference on finite populations. Density quantile estimation. Asymptotic theory of empirical processes, quantiles, functions of marginal quantiles, and L-statistics. Functional limits theorems. Large and Moderate deviations for dependent variables. Moderate deviations in general topological spaces. Probabilistic and Analytic number theory.

**Research Grants:**

2012–14 Principal Investigator for '2012 Summer School in Statistics for Astronomers'. $24,269. 3/15/2012 to 2/28/2014. NSF No. AST-1212302.

2011–13 Principal Investigator for '2011 Summer School in Statistics for Astronomers, and The Fifth Statistical Challenges in Modern Astronomy Conference'. $34,799. 5/1/2011 to 4/30/2013. NSF No. AST-1113001.

2010–13 Principal Investigator for 'SI2-SSE: Statistical software for astronomical surveys'. $450,000. 9/15/2010 - 8/31/2013. NSF No. AST-1047586.

2010–12 Principal Investigator for '2010 Summer School in Statistics for Astronomers'. $32,172. 3/1/2010 to 2/29/2012. NSF No. AST-1019605.

2009–11 Principal Investigator for '2009 Summer School in Statistics for Astronomers; June 8-13, 2009'. $33,359. 3/15/2009 to 2/28/2011. NSF No. AST-0915069.

2008–10 Principal Investigator for '2008 Summer School in Statistics for Astronomers; June 9-14, 2008', Co-PI Eric Feigelson. $33,131. 3/1/2008 to 2/28/2010. NSF No. AST-0808877.

2007–10 Principal Investigator for 'MSPA-AST: Advancing statistical methodology in massive astronomical surveys', Co-PI Eric Feigelson. $100,000. 9/15/2007 to 8/31/2010. NSF No. AST-0707833.

2007–08 Scientific Computing Research Environments for the Mathematical Sciences (SCREMS), Co-P.I. - G. J. Babu. $50,000. 9/1/2007 to 8/31/2008. NSF No. DMS-0722351.

2005–2006 Principal Investigator for 'Statistical Challenges in Modern Astronomy IV'. $20,000. 9/15/2005 to 12/14/2006. NASA grant No: NNG05GQ16G.

2005–2007 Principal Investigator for 'Research Experience for Undergraduates' supplement (NSF Grant No. AST-0535454) to "Astrostatistics: Advancing statistical methodology for Astronomy" $12,000. 7/5/2004-9/30/2007. NSF No: AST-0434234.

2004–2008 Principal Investigator for "Astrostatistics: Advancing statistical methodology for Astronomy", Co-PI Eric Feigelson. $508,359. 10/1/2004 to 9/30/2008. NSF No: AST-0434234.

2004–2007 Principal Investigator for "Center for Astrostatistics", $35,000. Penn State's Outreach Program Innovation Fund.

2003–2008 ITR: Grid Service Workflow System as a Research Environment for Science with Massive Data Sets, Co-I G. J. Babu (PI. Roy D. Williams) NSF No. AST-0326524. $3,117,508. (Penn State's share $215,000). 10/1/2003 to 3/31/2008.

2003–04 Scientific Computing Research Environments for the Mathematical Sciences (SCREMS), Co-P.I. - G. J. Babu. $65,347. 8/1/2003 to 7/31/2004. NSF No. DMS-0322673.

2003–04 Principal Investigator for 'Research Experience for Undergraduates' supplement (Grant No. DMS-0332264) to, "Multivariate Statistical Methodology for the Virtual Observatory". $10,000. 6/15/2001 to 5/31/2004. NSF No. DMS-0101360.

2001–04 Principal Investigator for the project, "Multivariate Statistical Methodology for the Virtual Observatory", Co-PI Eric Feigelson. $1,016,289. 6/15/2001 to 5/31/2004. NSF No. DMS-0101360.

2001–02 Principal Investigator for the project, "Conference on Statistical Challenges in Modern Astronomy III", Co-PI Eric Feigelson. $20,000. 3/1/2001 to 2/28/2002. NSF No. DMS-0096490.

2001–02 Principal Investigator for the project, "Conference on Statistical Challenges in Modern Astronomy III", funded by NASA. $20,000. 3/15/2001 to 3/14/2002. No. NAG5-10542.

2001 Principal Investigator for the project, "Conference on Statistical Challenges in Modern Astronomy III", funded by the Division of Continuing Education, Penn State University. (Program Innovation Fund) $4,000.

2000–01 Co-Principal Investigator for the project, "Conference on Statistical Challenges in Modern Astronomy III". $2,000. 10/1/2000 to 12/31/2001. The Pennsylvania Space Grant Consortium.

2000-01 Scientific Computing Research Environments for the Mathematical Sciences (SCREMS), Co-P.I. - G. J. Babu. $25,251. 9/1/2000 to 8/31/2001. NSF No. DMS-0079656.

1997–99 National Research Council's Twinning Fellowship, for travel to Lithuania and to host Professor E. Manstavičius for a month each year. $13,600. 9/1/1997 to 12/31/1999.

1997–99 Principal Investigator for the project, "Large Sample Methods". $45,503. 2/25/97 to 2/24/99. NSA No. MDA904-97-1-0023.

1996–99 Principal Investigator for the project, "Multivariate Estimation for Astronomy." $95,000. 8/1/96 to 7/31/99. NSF No. DMS-9626189.

1996–97 Co-principal Investigator for IAU Technical Workshop, "Conference on Statistical Challenges in Modern Astronomy II", funded by IAU. $1,831.95. From 4/1/96 till expended.

1995–97 Co-principal Investigator for the project, "Conference on Statistical Challenges in Modern Astronomy II", funded by NSF. $12,000. 6/15/95 to 5/31/97. No. DMS-9504783.

1995–96 Co-principal Investigator for the project, "Conference on Statistical Challenges in Modern Astronomy II", funded by Continuing and Distance Education, Penn State University. (Program Development Fund) $7,500.

1995–96 Principal Investigator for the project, "Conference on Statistical Challenges in Modern Astronomy II", funded by International Science Foundation, New York. Supported travel for two Russians.

1995–96 Principal Investigator for the project, "Conference on Statistical Challenges in Modern Astronomy II", funded by NASA. $15,000. 7/1/95 TO 12/31/96. No. NAGW-4793.

1994–95 Principal Investigator for the project "Planning Visit to India", funded by NSF. $2,200. January 1995 to December 1995. No. INT-9419424.

1992–95 Co-principal Investigator for the project, "Multivariate and Censored Data Analysis Methods for Astronomy", funded by NSF. 7-15-92 to 8-14-95. No. DMS-9208066.

1991 Principal Investigator for the project, "International Conference on – Statistical Challenges in Modern Astronomy", funded by NSF. $16,000. 1/1/91 to 12/31/91. No. DMS-9003083.

1990–1995 Co-Investigator for the project, "Multiwavelength and Statistical Research in Space Astrophysics", funded by NASA Long-term Space Astrophysics Research Program. 6/1/90 to 5/31/95. $120,000 per year. No. NAGW-2120.

1990–1992 Co-Principal Investigator for the project, "Advanced Statistical Methods for Analyzing Data from Astronomical Surveys", funded by NSF. 7/1/90 to 6/30/92. $68,450. No. DMS-9007717.

1990–1992 Principal Investigator for the project, "Edgeworth Expansions and Bootstrap", funded by National Security Agency. $30,000. 7/1/90 to 6/30/92. No. MDA904-90-H-1001.

1990–1992 Co-Principal Investigator for the project, "Advanced Statistical Methods for Improved Data Analysis of NASA Astrophysics Missions", funded by NASA's Astrophysics division. $83,000. 2/1/90 to 1/31/1992. No. NAGW-1917.

1989–1990 Co-Principal Investigator for the project, "Mathematical Sciences Research Equipment," funded by NSF. $41,683. July 1989 to June 1990. No. DMS–8905785.

1988–1990 Research Associate for the project, "Applications of advanced statistical methods to satellite survey," funded by Jet Propulsion Lab/NASA. $27,000. No. JPL-958013.

1988–1989 Co-Principal Investigator for the project, "Evaluate Statistical designs for travel surveys and formulate pilot sampling design", funded by Chesapeake Bay Stock Assessment Program of USDC. $75,000. 8/1/88 to 7/31/89.

1987–1988 Co-Principal Investigator for the project, "Mathematical Statistics and Statistics for Analysis and interpretation in Marine Fisheries Research and Management", funded by USDC. $73,333. 3/1/87 to 8/31/88.

1987–1988 Research Director and Principal Investigator for a project funded by Chesapeake Bay Stock Assessment Program. $50,000.

1986–1987 Co-Principal investigator, for two projects funded by NOAA. The total grant for these two projects is $195,000.

1985–1986 Worked with the Ecology group on a grant from National Oceanic and Atmospheric Administration (NOAA).

**Doctoral Thesis Advisor to:**

Kesar Singh, 1980, Indian Statistical Institute. Professor at Rutgers University, USA.
Bhaskar Bagchi, 1983, Indian Statistical Institute. Professor at Indian Statistical Institute, Bangalore, India.
Arup Bose, 1987, Indian Statistical Institute. Professor at Indian Statistical Institute, Calcutta, India.
Regis Serinko, 1990, The Pennsylvania State University.
Kang, Hee-Jeong, 1995, Chonbuk National University, Chonju, Korea.
Mark Leeds, 2000.
James McDermott, 2003, State Street Associates, Cambridge, MA.
Hyun-sook Lee, 2006, Harvard-Smithsonian, Center for Astrophysics, Cambridge, MA. Scott Roths, 2011, The Pennyslvania State University.

**Served on Doctoral Committees:**

1986–87 Ernst Linder (Statistics) – Special Member.
1986–87 Ken Suman (Statistics) – Special Member.
1988–89 Ranga V. Ramasesh (Business Administration).
1988–89 Joe Scazzero (Business Administration).
1988–89 Gunnar Stefansson (Mathematics).
1988–90 Zhijun Liu (Statistics).
1988–90 Min Deng (Statistics).
1989–90 Regis Serinko (Statistics) – Chairperson.
1989–91 Clint Coakley (Statistics).
1988–91 K. V. K. Prasad (Mining Engineering).
1989–92 Venkateshwar Reddy (Finance).
1989–92 Nandini Kannan (Statistics).
1991–95 Hee-Jeong, 1995 (Statistics) – Chairperson.
1996–97 Rajiv Dama (Mechanical Engineering).
1996–99 Arthur Dryver (Statistics).
1995–00 Mary E. Stocken (Business Administration).
1998–00 Mark Leeds (Statistics) – Co-chair.
2000–03 James McDermott (Statistics) – Co-advisor.

2000–05 Srikant Vadali (Business Administration).
2003–06 Hyun-sook Lee (Statistics) – Chairperson.
2005–07 Derek Young (Statistics).
2007–09 Kagan Kursungoz (Mathematics).
2007–11 Scott Roths (Statistics) – Chairperson.
2007–11 Arseny Egorov (Mathematics).
2009– Day Prapanpong (Mathematics).

**Masters Theses Supervised:** Regis Serinko (1988-1989), Hyun-sook Lee (2001-2003).

**Served on Master's Thesis Committees:** Anamagdalen Nitica (1994), Derek Young (2005).

**Undergraduate Students:** (*Research Experience for Undergraduates*) Xiao-Yi Li (BS Honors) 2003, Stas Sheynkop (BS Honors) 2004, Tae W. Kang 2004, Michael L. Rogers, 2005, and Matthew A. Lohr (BS Honors) 2006.

**College and Departmental Committees:**

Chair, Selection Committee for the John M. Chemerda Lectures in Science, 2003–2004.

Member, Eberly College of Science Strategic Vision Committee, 2004–2005.

Member, Eberly College of Science Outreach Council, 2005–Present.

Member, Eberly College of Science Sabbatical Leave Review Committee, 2006.

Selection Committee for the John M. Chemerda Lectures in Science, 1988–1989, 1992–1993.

Eberly College of Science Promotion and Tenure Committee, 1993–1995, 2000–2002.

College of Science Promotion and Tenure Committees:
*Mathematical Sciences* (1988, 1989, 1991), and *Biological Sciences* (1990).

Eberly College of Science Immediate Tenure committee, 1998, 2004.

Professor-in-charge, Master of Applied Statistics Program, 2001–2007.

Chair, Department of Statistics Promotion and Tenure Committee, 1989.

Chair, Statistics Department's World Campus Program, 1998–2002.

Chair, C.R. and Bhargavi Rao Prize selection committee, 2003, 2005, 2007, 2009, 2011.

Chair, Awards Committee, Department of Statistics, 2004–2005.

Chair, Fifth year post tenure evaluation committee, Department of Statistics, 2000, 2006, 2010.

Chair, Ph.D. qualifier exam committee, Department of Statistics, 2010 (Member 2009).

Faculty Recruiting Committee, Department of Statistics, 2001–2004.

Graduate program committee, Department of Statistics, since 2005.

Interdisciplinary Activities Committee, 2008-2009.

Fixed term faculty review committe, 2010.

Chair, Faculty hiring committee, Department of Statistics, 1996–1997 and 2011–2012.

**University Committees:**

Chair, Committee on Academic Standards, Graduate Council, Graduate Council (1997–1998, 2004–2005).

Chair, Subcommittee on World Campus (Faculty Senate committee on outreach activities 1999-2000).

Committee on Academic Standards, Graduate Council (1997–1999, 2003–2005).

Elected to Graduate Council (2012-2014).

Alternate on Graduate Council (2011-2012).

Executive Committee, Graduate Council (1998-1999, 2003–2005).

Ad Hoc Committee on Professional Master's Programs, Graduate Council (2004–2005).

Committee on Committees and Procedures, Graduate Council (1997–1999).

Penn State Faculty Outreach Award Selection Committee (1998, 2000-2002).

Implementation Committee on Electronic Theses and Dissertations 1998–2003.

ETD sub-committee on Communication and Information (1999).

Penn State World Campus Steering Committee (1998–1999, 2001–2005).

Graduate Council's World Campus Working Group (1999).

Faculty Senate Committee on Outreach Activities (1997–2002).

Graduate Council Subcommittee on Programs Review and Evaluation (2000–2001).

Penn State Review Panel on Outreach Scholarship Conference (a partnership between Penn State, the University of Wisconsin-Extension, Ohio State, and the University of Georgia, 2006).

**Organization of Conferences and Schools:**

*Co-chair*, Scientific Organizing Committee for the international conference on *Statistical Challenges in Modern Astronomy*. August 11–14, 1991. This inter–disciplinary conference is cosponsored by *IMS, NASA* and *NSF*.

Organized an invited paper session on *Statistics in Astronomy*, at 1993 Joint Statistical Meetings in San Francisco. Also chaired the session.

*Co-chair*, Scientific Organizing Committee for the international conference on *Statistical Challenges in Modern Astronomy II*. June 2–5, 1996. This inter–disciplinary conference is cosponsored by *IMS, ISI, IAU, ISF, NASA* and *NSF*.

Organized an invited paper session on 'Astrostatistics' for International Astronomical Union at 51st session of the International Statistical Institute, August 18-27, 1997.

Organized a session on 'Astrostatistics' (Track: Emerging Science: Transforming the Next Generation) at AAAS Annual meeting and science innovation exposition in Philadelphia, in February 1998.

Organized 'Subramanyan Chandrasekhar Memorial Session on Astrostatistics' on August 12, 1999, Joint Statistical Meetings in Baltimore.

Organized a session on 'Data Mining' in August 2000, at Joint Statistical Meetings in Indianapolis.

Member, Organizing Committee for the International Conference in Statistics in Calcutta, India during December 29-30, 2000.

Organized a session on 'Resampling methods' at the International Conference in Statistics in Calcutta, India during December 29-30, 2000.

Member, Advisory Committee for Joint Statistical Meeting in New Delhi, India during December 30, 2000 - January 2, 2001.

Organized a session on 'Astrostatistics' at the Joint Statistical Meeting in New Delhi, India during December 30, 2000 - January 2, 2001.

*Co-chair*, Scientific Organizing Committee for the international conference on *Statistical Challenges in Modern Astronomy III*. The Pennsylvania State University, July 18-21, 2001. This inter–disciplinary conference is cosponsored by *NASA, Penn State University* and *Pennsylvania Space Grant Consortium*.

Member, Program Committee for the conference 'Astronomical Data Analysis', part of *The International Society for Optical Engineering*'s (SPIE) Symposium on Optical Science and Technology, in San Diego, 29 July - 3 August 2001.

Member, Program Committee for the conference 'Astronomical Data Analysis II', part of *SPIE*'s Symposium on Astronomical Telescopes and Instrumentation, in Waikoloa, Hawaii, 22 - 28 August 2002.

Member, Scientific Organizing Committee at for the conference *Astronomical Data Analysis- III*, in Sant' Agata sui due Golfi (NA), Italy, 29 April to 1 May 2004.

Organized a session on 'Astrostatistics' at the International Conference on the *Future of statistical theory, practice and education*, at Hyderabad, India. December 29, 2004 - Jan 1, 2005.

Organized a *Summer School in Statistics for Astronomers and Physicists* at Penn State, during June 5-17, 2005.

Organized a *SAMSI Astrostatistics program planning meeting* at NASA Ames Center during July 14-15, 2005.

Chair of the Program Leaders Committee, and directed a semester long *Astrostatistics program* at SAMSI during January - May 2006.

Organized *Tutorials* for astronomers and statisticians at SAMSI (January 18-22, 2006).

Organized the *Opening Workshop* to focus on the scientific agenda of the Astrostatistics program at SAMSI (January 23-25, 2006).

Organized the second *Summer School in Statistics for Astronomers and Physicists* at Penn State, during June 6-10, 2006.

Co-Chair, Scientific Organizing Committee for 'Statistical Challenges in Modern Astronomy IV', held at Penn State University during June 12-15, 2006.

Member, Scientific Organizing Committee for *Astronomical Data Analysis IV*, held at Laboratoire d'Astrophysique deMarseille, Marseille, France, during September 18-20, 2006.

Organized the third *Summer School in Statistics for Astronomers* at Penn State, during June 4-9, 2007.

Organized a *Summer School in Statistics for Astronomers*, in collaboration with *The Indian Institute of Astrophysics, Bangalore*, at the Vainu Bappu Observatory near the village of Kavalur in India, during July 2-7, 2007.

Member, Scientific Organizing Committee for *Astronomical Data Analysis V*, Heraklion, Crete (Greece), during May 7-9, 2008.

Organized the fourth *Summer School in Statistics for Astronomers* at Penn State, during June 9-14, 2008.

Organized the second *Summer School in Statistics for Astronomers*, in collaboration with *The Indian Institute of Astrophysics, Bangalore*, at the Vainu Bappu Observatory near the village of Kavalur in India, during July 9-16, 2008.

Organized the fifth *Summer School in Statistics for Astronomers* at Penn State, during June 1-6, 2009.

Member, Scientific Organizing Committee for *III INPE Advanced Course on Astrophysics: Astrostatistics*, to be held at Sao Jose dos Campos-SP, Brazil, during September 14-18, 2009.

Member, Scientific Organizing Committee for *ADA VI – Astronomical Data Analysis*, conference in Honor of Albert Bijaoui, to be held at Monastir, Tunisia, during May 3-6, 2010.

Organized the sixth *Summer School in Statistics for Astronomers* at Penn State, during June 7-12, 2010. This was followed by a supplementary program on Statistics and Computation for Astronomical Surveys (June 12-14, 2010).

Organized the third *Summer School in Statistics for Astronomers*, in collaboration with *The Indian Institute of Astrophysics, Bangalore*, at the Vainu Bappu Observatory near the village of Kavalur in India, during July 19-27, 2010.

Organized the seventh *Summer School in Statistics for Astronomers* at Penn State, during June 6-10, 2011. This was followed by a two-day tutorials on three different topics (June 11-12, 2011).

Co-Chair, Scientific Organizing Committee for the cross-disciplinary conference, 'Statistical Challenges in Modern Astronomy V' to be held in June 13-17, 2011.

Organized an astrostatistics school for the Instruments Division at the Space Telescope Science Institute (STScI, http://www.stsci.edu/portal/) in September - November 2011.

Member, Scientific Organizing Committee, *ADA VII – Astronomical Data Analysis* conference, Cargese, Corsica, France, May 14-18, 2012.

**Invited Conferences:**

On a grant from International Mathematical Union, attended the *International Congress of Mathematicians* at Helsinki, Finland in August 1978.

Oberwolfach Conference on *Analytic Number Theory* in November 1980 in West Germany.

Fourth Annual Conference of the *Indian Society for the Theory of Probability and its Applications* held at Indian Institute of Management, Calcutta in June 1983.

Fourth Mathscience Conference on *Number Theory*, held at Ooty, India. December, 1984.

Special Symposium on U. S. National *Monitoring Strategies*. Oceans 86. Washington D.C., during September 23–25, 1986.

ASA and EPA Conference on *Statistical Issues in Combining Environmental Studies*, at Washington D.C., during October 1-2, 1986.

AMS/SIAM/IMS Summer Research Conference on *Statistical Analysis of Measurement Error Models and Applications*, held at Humboldt State University during June 10–16, 1989.

International Conference on Recent Developments in *Statistical Data Analysis and Inference* held at the University of Neuchâtal, Switzerland, during August 21–24, 1989. *Chaired a Session.*

*Statistics '91 Canada,* Third Canadian Conference in Applied Statistics held at Concordia University, Montreal, Canada. May 23 - May 25, 1991.

International conference on *Analytic and Probabilistic Methods in Number Theory* at Palanga, Lithuania, USSR. September 24-28, 1991. (Presented an Invited paper, could not attend due to political unrest in USSR.)

Some Recent Contributions to Edgeworth Expansions. An invited talk presented at *Probability Day* at Penn State University. July 15, 1991.

First International Triennial Calcutta Symposium on *Probability and Statistics* at Calcutta University, India. December 27, 1991 to January 1, 1992.

Statistics in Astronomy. (With Eric Feigelson). Invited talk presented at 1993 Joint Statistical Meetings in San Francisco on August 11, 1993.

Bootstrap and other Resampling Methods. An invited talk presented at *Bootstrap Day* at Penn State University. November, 12, 1994.

Bootstrap and other resampling schemes. A 90 minute invited talk on given at *NSF Summer Symposium on the Bootstrap 1996* at Econometrics Laboratory, University of California at Berkeley. July 30 - August 6, 1996.

Conjecture by Erdős and additive functions on the set of pairs of integers. Invited talk given on September 24, 1996 at *II International Conference on Analytic and Probabilistic Number Theory* organized by Vilnius University during September 23-27, 1996 at Palanga, Lithuania, to honor Professor Jonas Kubilius on his 75th birthday.

The resurgence of astrostatistics. Invited talk at *51st session of the International Statistical Institute* at Istanbul, 18-26 August 1997. The session on Astrostatistics is sponsored by International Astronomical Union.

Conditional Edgeworth Expansions and Resampling Schemes. Invited talk at *85th Session of the Indian Science Congress*, Osmania University, Hyderabad, India, January 3-7, 1998.

Asymptotic theory for random permutations with applications to genetics. Invited talk at 1998 Lukacs Symposium, *Statistics for the 21st Century*, Bowling Green University, April 24-26, 1998.

Comparison of Resampling Procedures. Invited talk at *Rutgers Bootstrap Conference*, May 14-16, 1998.

Breakdown theory for estimators based on bootstrap and other resampling schemes. Invited talk at *IISA International conference 1998*, McMaster University, Hamilton, Canada, October 10-11, 1998.

Random Permutations and the Ewens sampling formula in genetics, at the International conference on *Combinatorics, Statistics, Pattern Recognition and Related Areas*, Mysore, India, on December 29, 1998.

Limit Theorems for Random Permutations, at the conference, *Paul Erdős and his Mathematics*, Budapest, Hungary July 4-11, 1999.

Functional Limit Theory for processes generated by random permutations. Invited talk at *Workshop on the Interface of Probability and Number Theory*, University of Illinois at Urbana-Champaign, May 19-20, 2000.

Statistical methodology for NVO. Invited talk at the *Conference on Virtual Observatories of the Future*, Caltech, June 13 - 16, 2000.

Statistical Methodology for NVO. ESO/ESA/NASA/NSF Astronomy Conference, *Toward an International Virtual Observatory*, June 10 - 14, 2002, Garching, Germany.

Goodness-of-fit tests when parameters are estimated. Special invited lecture at the *Fourth Biennial International Conference on Statistics, Probability and Related Areas*, June 14-16, 2002, Northern Illinois University in DeKalb, Illinois.

Statistical methodology for massive datasets and model selection, *Astronomical Data Analysis II (ADA)*, part of SPIE Symposium on Astronomical Telescopes and Instrumentation, in Waikoloa, Hawaii, 22-28 August 2002.

Invited participant at the 'Workshop on Statistical methods for the analysis of massive streams of data', organized by the National Academies' Committee on Applied & Theoretical Statistics. National Academies Building, Washington, D.C. December 13-15, 2002.

Statistical and computational challenges, and opportunities in Astronomy. Invited talk at *International Conference on Ranking and Selection, Multiple Comparisons and Reliability, and Their Applications*, December 28-30, 2002, Chennai, India

Goodness of fit tests with estimated parameters. Invited talk at *XXIII International Seminar on Stability Problems for Stochastic Models*, May 12-17, 2003, Pamplona, Spain.

Statistical challenges in Modern Astronomy (with E. Feigelson). Invited presentation at *International Conference on Advanced Statistical Methods in Particle and Astro-Particle Physics*, September 8-12, 2003, SLAC, Stanford, CA.

Probabilistic Number Theory and Random Permutations: Functional Limit Theory. Invited presentation at *Conference on Zeta Function*, December 13-15, 2003, National Institute of Advanced Studies (NIAS), Bangalore, India.

Model fitting in the presence of nuisance parameters. Invited presentation at the conference *Astronomical Data Analysis-III*, Sant' Agata sui due Golfi (NA), Italy, 29 April to 1 May 2004.

Invited participant at the *Statistical Issues in Data Acquisition*, organized by the National Academies' Committee on Applied & Theoretical Statistics. Lawrence Berkeley National

Laboratory, Berkeley, July 16, 2004. http://sites.nationalacademies.org/DEPS/BMSA/DEPS_047678.

Statistical Challenges in Modern Astronomy. Invited presentation at International Conference on *Recent developments in statistics and their applications*, Tirupati, India. January 3-4, 2005.

Invited participant at *Statistics for Gravitational Wave Data analysis* workshop, Penn State. May 19-21, 2005.

Weak convergence for additive functions on random partitions of an integer. Invited speaker at the international conference on *Probability and Number partitions Theory*, Kanazawa, Japan. June 20-24, 2005.

Goodness-of-fit and all that! Invited speaker at *ADASS IV - Astronomical Data Analysis Software & Systems XV*, San Lorenzo de El Escorial, Spain. October 2-5, 2005.
Tutorial. R: A powerful public software environment for statistical analysis of astronomical data. (with David Hunter, Eric Feigelson). 170 participants

Statistical Problems in Astronomy (December 21, 2005).
Resampling Techniques (December 22, 2005). Invited speaker at the *Workshop on Astrostatistics*, Calcutta University, Kolkata, India, December 21-23, 2005.

Invited Discussant at *Statistical Challenges in Modern Astronomy IV*, Penn State University, June 12-15, 2006

Object detection in multi-epoch data. Invited speaker at *Astronomical Data Analysis IV*, Laboratoire d'Astrophysique deMarseille, Marseille, France. September 18-20 2006.

Invited participant at the *Phystat-LHC Workshop on Statistical issues for LHC Physics*, CERN, Geneva, June 27-29, 2007.

Marginal quantiles: Asymptotics for functions of order statistics. Invited speaker at the conference on *Recent Advances in Probability* held at the Indian Statistical Institute, Calcutta. December 11-15, 2007. Part of Platinum Jubilee celebration of the Indian Statistical Institute.

Invited participant at *The LSST All-Hands Meeting*, the National Center for Supercomputing Applications, University of Illinois, Urbana-Champain, May 19-23, 2008.

Edgeworth expansions and their applications. Invited speaker at the conference on *Advances in Statistics*, held in honor of the 65th birthday of Zhidong Bai at National University of Singapore, July 20, 2008.

Understanding 21st Century Astronomical Data Cubes. Invited speaker at the IJCAI-09 Workshop on *Machine Learning and AI Applications in Astrophysics and Cosmology*, Pasadena, California, July 16-17, 2009.

Participant at the 'Prague Stochastics 2010' conference held at Charles University in Prague, Czech Republic during August 30 to September 3, 2010.

Feature identification in datacubes. Invited speaker at *ALMA Software Development Workshop*, National Radio Astronomy Observatory, Charlottesville, VA, October 12-14, 2011.

Analysis of Astronomical data cubes. Invited speaker at *Digging Deeper: Algorithms for Computationally-Limited Searches in Astronomy - Part II* at Caltech in Pasadena, California., sponsored by

the Keck Institute for Space Studies, December 12-15, 2011.

Invited panelist at US/India NSF Workshop on *Virtual Institutes for Computational and Data-Enabled Science & Engineering*, Bangalore, India on December 21-22, 2011. Panelist for Data-Intensive Computing & Astrophysics.

**Invited Talks:**

International institutions, where invited talks are given include: University of Oslo (Norway), University of Lund (Sweden), University of Copenhagen (Denmark), Mathematical Institute of Paderborn (West Germany), University of Szeged (Hungary), Andhra University (India), Monash University (Australia), Melbourne University (Australia), Latrobe University (Australia), Australian National University (Australia), Indian Statistical Institute, (Kolkata, India), Keio University (Japan), Institute of Statistical Mathematics (Japan), Hiroshima University (Japan), Osaka city University (Japan), National Sun Yet-sen University (Taiwan), Concordia University (Canada), University of Hyderabad (India), University of Poona (India), Inter-University Centre for Astronomy and Astrophysics (India), National University of Singapore (Singapore), National Tsing Hua University (Taiwan), Institute of Statistical Science, Academia Sinica (Taiwan), LAMOST project of National Astronomical Observatory of the Chinese Academy of Sciences (Beijing, China), Northeast Normal University (Changchun, China), Indian Institute of Astrophysics (Bangalore, India), Raman Research Institute (Bangalore, India), and other universities and institutes.

Institutions in the USA, where invited colloquia talks are given include: Rutgers University, University of Pennsylvania, University of Maryland at Baltimore County, Michigan State University, Case Western Reserve University, Mathematical Sciences Research Institute (Berkeley), University of Georgia, Florida State University, Purdue University, Iowa State University, Ohio State University, Wright State University, Texas A&M university, Statistical and Applied Mathematical Sciences Institute (Research Triangle Park, North Carolina), North Carolina State University, University of South Carolina, University of Minnesota, University of California at Irvine, The Pennsylvania State University (Department of Industrial and Manufacturing Engineering), University of Texas at Dallas, Columbia University (New York), Temple University, University of California at San Diego, Univesity of California at Berkeley, and other institutes and universities.

# List of Research Publications of G. J. Babu

The numbers following **MR** denote the *Mathematical Reviews* numbers.

**Books Published:**

1. Feigelson, E. D. and Babu, G. J. (2012). *Modern Statistical Methods for Astronomy with R applications*. Cambridge University Press, Cambridge. In Press.

2. Babu, G. J., and Feigelson, E. D. (Editors) (2007). *Statistical Challenges in Modern Astronomy IV*. ASP Conference Series, Vol. 371, Astronomical Society of the Pacific, San Francisco.

3. Feigelson, E. D., and Babu, G. J. (Editors) (2003). *Statistical Challenges in Astronomy*. Springer-Verlag, New York.

4. Babu, G. J., and Feigelson, E. D. (Editors) (1997). *Statistical Challenges in Modern Astronomy II*. Springer-Verlag, New York.

5. Babu, G. J., and Feigelson, E. D. (1996). *Astrostatistics*. Chapman and Hall, London. Reprinted (1997).

6. Feigelson, E. D., and Babu, G. J. (Editors) (1992). *Statistical Challenges in Modern Astronomy*. Springer-Verlag, New York.

7. Babu, G. J. (1978). *Probabilistic Methods in the Theory of Arithmetic Functions.* Macmillan Lecture Series, Series 2, New Delhi. (**MR** #80g: 10057).

**Research Publications:**

1. Lee, H., Babu, G. J., and Rao, C. R. (2012). A jackknife type approach to statistical model selection. *Journal of Statistical Planning and Inference*, **142**, issue 1, 301311.

2. Babu, G. J. Resampling methods for model fitting and model selection. (2011). *Journal of Biopharmaceutical Statistics*, **21**, issue 6, 1177-1186.

3. Feigelson, E. D., and Babu, G. J. Statistical Methods for Astronomy. To appear in *Planets, Stars and Stellar Systems*, Edited by Terry Oswalt. Springer, New York, NY.

4. Babu, G. J., Bai, Z. D., Choi, K.-P., and Mangalam, V. (2011). Limit Theorems for functions of marginal quantiles. *Bernoulli*, **17**, Number 2, 671-686.

5. Babu, G. J., Chattopadhyay, T., Chattopadhyay, A., and Mondal. S. (2009). Horizontal branch morphology of globular clusters: A multivariate statistical analysis. *The Astrophysical Journal*, **700**, 1768-1778.

6. Babu, G. J. (2009). Marginal quantiles: Asymptotics for functions of order statistics. In *Perspectives in Mathematical Sciences I: Probability and Statistics*. Statistical Science and Interdisciplinary Research, **Vol. 7**. World Scientific Publishing Co., 31-39.

7. Babu, G. J., Mahabal, A., Williams, R., and Djorgovski, S. G. (2008). Object detection in multi-epoch data. *Statistical Methodology*, **5**, issue 4, 299-306. [Slides]

8. Babu, G. J. (2008). Edgeworth Expansions: A brief review of Zhidong Bai's contributions. In *ADVANCES IN STATISTICS*, Zehua Chen, Jin-Ting Zhang & Feifang Hu (Eds.). World Scientific Publishing Co, 16-18.

9. Babu, G. J., and Padmanabhan, A. R. (2007). Re-sampling methods for testing for location against unrestricted and ordered alternatives. *Journal of Statistical Planning and Inference*, **137**, issue 11, 3261-3267.

10. Babu, G. J., Manstavičius, E., and Zacharovas, V. (2007). Limiting processes with dependent increments for measures on symmetric group of permutations. *Advanced Studies in Pure Mathematics*, **49**, 41-67. (**MR** #2009h: 60022).

11. McDermott, J. P., Babu, G. J., Liechty, J. C., and Lin, Dennis K. J. (2007). Data skeletons: simultaneous estimation of multiple quantiles for massive streaming datasets with applications to density estimation. *Statistics and Computing*, **17**, issue 4, 311-321.

12. Babu, G. J., and Mahabal, A. (2007). Using R-based VOStat as a low-resolution spectrum analysis tool. *Journal of Statistical Software*, **18**, issue 11, 1-12.

13. Babu, G. J. (2006). Probabilistic Number Theory and Random Permutations: Functional Limit Theory. In *The Riemann Zeta function and related themes*, R. Balasubramanian and K. Srinivas (eds.). Ramanujam Mathematical Society – Lecture Notes Series, No. 2, 19-27. (**MR** #2008g: 60019).

14. Babu, G. J., and Chaubey, Y. P. (2006). Smooth estimation of a distribution and density function on hypercube using Bernstein polynomials for dependent random vectors. *Statistics and Probability Letters*, **76**, no. 9, 959-969. (**MR** #2008h: 62078).

15. Babu, G. J., and Djorgovski, S. G. (2004). Some statistical and computational challenges, and opportunities in astronomy. *Statistical Science*, **19**, no. 2, 322-332.

16. Babu, G. J., and Rao, C. R. (2004). Goodness-of-fit tests when parameters are estimated. *Sankhyā*, **66**, no. 1, 63-74. (**MR** #2005c: 62053).

17. Babu, G. J. (2004). A note on the bootstrapped empirical process. *J. of Statistical Planning and Inference*, **126**, no. 2, 587-589. (**MR** #2006a: 62053).

18. Babu, G. J., Boyarsky, A., Chaubey, Y. P. and Gora, P. (2004) New statistical method for filtering and entropy estimation of a chaotic map from noisy data. *International Journal of Bifurcation and Chaos*, **14**, no. 11, 3989-3994. (**MR** #2005j: 37136).

19. Babu, Gutti Jogesh, and Rao, M. Bhaskara. (2004). Occurrence/exposure rate. In *Encyclopedia of Actuarial Science*, Bjoern Sundt and Jef Teugels (Eds.), Wiley, Chichester, Vol **3**, 1199-1201.

20. Babu, G. J., and Rao, C. R. (2003). Confidence limits to the distance of the true distribution from a misspecified family by bootstrap. *J. Statistical Planning and Inference*, **115**, no. 2, 471-478. (**MR** #2004c: 62072).

21. Scargle, J. D., and Babu, G. J. (2003). Point processes in astronomy: Exciting events in the universe. *Handbook of Statistics,* **Vol. 21** "Stochastic Processes: Modeling and Simulation." C. R. Rao and D. N. Shanbhag (Eds.), Elsevier Science Publishers B. V., Amsterdam, 795-825.

22. Babu, G. J., Singh, K., and Yang, Y. (2003). Edgeworth expansions for compound poisson processes and the bootstrap. *The Annals of the Institute of Statistical Mathematics*, **55**, no. 1, 83-94. (**MR** #2004b: 62084).

23. Babu, G. J., and Padmanabhan, A. R. (2002). Re-sampling methods for the nonparametric Behrens-Fisher problem. *Sankhyā*, Series A, **64**, 678-692.

24. Babu, G. J., and Manstavičius, E. (2002). Infinitely divisible limit processes for the Ewens sampling formula (Russian). *Lietuvos Matematikos Rinkinys*, **42**, no. 3, 294-307. English translation in *Lithuanian Math. J.*, **42**, no. 3, (2002), 232-242. (**MR** #2003k: 60044).

25. Babu, G. J., and Manstavičius, E. (2002). Limit processes with independent increments for the Ewens sampling formula. *The Annals of the Institute of Statistical Mathematics*, **54**, no. 3, 607-620. (**MR** #2003j: 60041).

26. Babu, G. J., Canty, A., and Chaubey, Y. (2002). Application of Bernstein Polynomials for Smooth Estimation of a Distribution and Density Function. *J. Statistical Planning and Inference*, **105**, no. 2, 377-392. (**MR** #2003d: 62088).

27. Babu, G. J., Pathak, P. K., and Rao, C. R. (2000). Consistency and accuracy of the sequential bootstrap. In "Statistics for the 21st century: Methodologies for application of the Future." C. R. Rao and G. Szekély (Eds.), Marcel Dekker, Inc, New York, 21-31.

28. Babu, G. J., Pathak, P. K., and Rao, C. R. (1999). Second order correctness of the Poisson bootstrap. *Annals of Statistics*, **27**, no. 5, 1666-1683. (**MR** #2001c: 62059).

29. Babu, G. J., and Manstavičius, E. (1999). Brownian motion for random permutations. *Sankhyā*, Series A, **61**, 312-327. (**MR** #2001j: 60016).

30. Babu, G. J. (1999). Breakdown theory for estimators based on bootstrap and other re-sampling schemes. In *Asymptotics, Nonparametrics, and Time Series.* Subir Ghosh (Ed.). Marcel Dekker, New York, 669-681.

31. Babu, G. J., Padmanabhan, A. R., and Puri, M. L. (1999). Robust one-way ANOVA under possibly non-regular conditions. *Biometrical Journal*, **41**, 321-339. (**MR** #2000d: 62108).

32. Babu, G. J., and Manstavičius, E. (1999). Random permutations and the Ewens sampling formula in genetics. *Probability and Mathematical Statistics.* B. Grigelionis *et al.* (Eds.), TEV, Vilnius and VSP, Utrecht, Netherlands, pp 33-42.

33. Mukherjee, S., Feigelson, E. D., Babu, G. J., Murtagh, F., Fraley, C., and Raftery, A. (1998). Three types of Gamma ray bursts. *The Astrophysical Journal*, **508**, November 20, 314-327.

34. Padmanabhan, A. R., Chinchilli, V. M., and Babu, G. J. (1997). Robust analysis of within-unit variances in repeated measurement experiments. *Biometrics*, **53**, 1520-1526.

35. Babu, G. J. (1997). On a conjecture by Erdős and its extension to additive functions on the set of pairs of integers. *New Trends in Probability and Statistics*, Vol 4. Analytic and Probabilistic Methods in Number Theory. A. Laurincikas, E. Manstavičius and V. Stakenas (Eds.); TEV, Vilnius and VSP, Utrecht, Netherlands, 261-270. (**MR** #2000a: 11136).

36. Babu, G. J. (1997). Bootstrap – A review. In *Probability and its applications.* M.C. Bhattacharjee and Sujit Basu (Eds.), The Oxford University Press, Delhi, 167-178.

37. Babu, G. J., and Bai, Z. D. (1996). Mixtures of global and local Edgeworth expansions and their applications. *J. Multivariate Analysis*, **59**, No. 2, 282-307. (**MR** #98d: 62023).

38. Babu, G. J., and Feigelson, E. D. (1996). Spatial point processes in astronomy. *J. Statistical Planning and Inference*, **50**, 311-326. (**MR** #97a: 85002).

39. Babu, G. J., and Chaubey, Y. P. (1996). Asymptotics and bootstrap for inverse gaussian regression. *Ann. Inst. Statist. Math.*, **48**, 75-88. (**MR** #97d: 62027).

40. Babu, G. J., and Padmanabhan, A. R. (1996). A robust test for omnibus alternatives. In "Research Developments in Probability and Statistics. Festscrift in honor of Madan L. Puri on the occasion of his 65th birthday" eds., E. Brunner and M. Denker. VSP International Science Publishers, 319-327. (**MR** #98f: 62033).

41. Koti, K., and Babu, G. J. (1996). Sign test for ranked-set sampling. *Communications in Statistics - Theory and Methods*, **25**, No. 7, 1617-1630.

42. Babu, G. J. (1995). Bootstrap for nonstandard cases. *J. Statistical Planning and Inference*, **43**, 197-203. (**MR** #96c: 62080).

43. Serinko, R. J., and Babu, G. J. (1995). Asymptotics of k-mean clustering under non-i.i.d. sampling. *Probability and Statistics Letters*, **24**, 57-66. (**MR** #96d: 62108).

44. Babu, G. J. (1995). Absolute continuity of the distributions of additive arithmetic functions. *Sankhyā*, Series A, **57**, 29-32. (**MR** #97d: 11144).

45. Linder, E., and Babu, G. J. (1994). Bootstrapping the linear functional relationship with known error variance ratio. *Scandinavian Journal of Statistics*, **21**, 21-39. (**MR** #95b: 62056).

46. Babu, G. J., and Bai, Z. D. (1993). Edgeworth expansions of a function of sample means under minimal moment conditions and partial Cramér's condition. *Sankhyā*, Series A, **55**, 244-258. (**MR** #95m: 60037).

47. Babu, G. J. (1993). Edgeworth expansions in non-regular cases and their applications to bootstrap. In "Statistics and Probability: A Raghu Raj Bahadur Festscrift" eds., J. K. Ghosh, S. K. Mitra, K. R. Parthasarathy and B. L. S. Prakasa Rao. Wiley Eastern Limited, New Delhi. 63-71.

48. Babu, G. J., and Rao, C. R. (1993). Bootstrap methodology. In *Handbook of Statistics,* **Vol. 9** "Computational Statistics." C. R. Rao (Ed.), Elsevier Science Publishers B. V., Amsterdam, 627-659. (**MR** #94j: 62097).

49. Babu, G. J., Rao, C. R., and Rao, M. B. (1992). Nonparametric estimation of specific occurrence/exposure rate in risk and survival analysis. *J. American Statistical Association*, **87**, 84-89. (**MR** #93k: 62080).

50. Babu, G. J. (1992). Subsample and half-sample methods. *Annals of the Institute of Statistical Mathematics*, **44**, 703-720. (**MR** #93k: 62101).

51. Babu, G. J., and Rao, C. R. (1992). Expansions for statistics involving the mean absolute deviations. *Annals of the Institute of Statistical Mathematics*, **44**, 387-403. (**MR** #93m: 62038).

52. Feigelson, E. D., and Babu, G. J. (1992). Linear regression in astronomy - II. *Astrophysical Journal,* **397**, September 20, 55-67.

53. Babu, G. J., and Bai, Z. D. (1992). Edgeworth expansions for errors-in-variables models. *J. Multivariate Analysis*, **42**, 226-244. (**MR** #93j: 62168).

54. Serinko, R. J., and Babu, G. J. (1992). Weak limit theorems for univariate k-mean clustering under a nonregular condition. *J. Multivariate Analysis*, **41**, 273-296. (**MR** #93k: 60062).

55. Babu, G. J. (1992). Smoothness of the distributions of arithmetic functions. *New Trends in Probability and Statistics*, Vol 2. Analytic and Probabilistic Methods in Number Theory. F. Schweiger and E. Manstavičius (Eds.); TEV, Vilnius and VSP, Utrecht, Netherlands, 191-199. (**MR** #93m: 11079).

56. Babu, G. J., and Feigelson, E. D. (1992). Analytical and Monte Carlo comparisons of six different linear least squares fits. *Communications in Statistics - Simulation and Computation,* **21** (2), 533-549. (**MR** #93e: 62167).

57. Bose, A., and Babu, G. J. (1991). Accuracy of the bootstrap approximation. *Probability theory and related fields*, **90**, 301-316. (**MR** #92m: 62050).

58. Babu, G. J. (1991). Asymptotic theory for estimators under random censorship. *Probability theory and related fields*, **90**, 275-290. (**MR** #92i: 62031).

59. Babu, G. J. (1991). Edgeworth expansions for statistics, which are functions of lattice and non-lattice variables. *Statistics and Probability letters*, **12**, 1-7. (**MR** #92k: 62027).

60. Rao, M. B., Babu, G. J., and Rao, C. R. (1991). Nonparametric estimation of survival functions under dependent competing risks. *Nonparametric Functional Estimation and Related Topics*, G. Roussas (Ed.), Kluwer Academic Publishers, 431-441. (**MR** #93a: 62063).

61. Babu, G. J. (1991). Discussion of P. K. Sen's paper, "Nonparametrics: Retrospectives and Perspectives." *Nonparametric Statistics*, **1**, 33-35.

62. Babu, G. J., and Rao, C. R. (1990). Estimation of the reciprocal of the density quantile function at a point. *J. Multivariate Analysis*, **33**, 106-124. (**MR** #91h: 62037).

63. Singh, K., and Babu, G. J. (1990). On asymptotic optimality of the bootstrap. *Scandinavian Journal of Statistics*, **17**, 1-9. (**MR** #91h: 62046).

64. Babu, G. J. (1990). A note on Dubins' theorem. *Canadian Mathematical Bulletin*, **33**, 416-418. (**MR** #92c: 60002).

65. Isobe, T., Feigelson, E. D., Akritas, M. G., and Babu, G. J. (1990). Linear regression in astronomy I. *The Astrophysical Journal*, **364**, November 20, 104-113.

66. Babu, G. J., and Singh, K. (1989). On Edgeworth expansions in the mixture cases. *Annals of Statistics*, **17**, 443-447. (**MR** #90b: 62019).

67. Babu, G. J., and Erdős, P. (1989). On the distribution function of additive arithmetical functions in short intervals. *Canadian Mathematical Bulletin*, **32**, 441-445. (**MR** #90i: 11106).

68. Babu, G. J. (1989). Strong representations for LAD estimators in linear models. *Probability Theory and Related Fields*, **83**, 547-558. (**MR** #90i: 62079).

69. Babu, G. J., and Singh, K. (1989). A note on Edgeworth expansions for the lattice case. *J. Multivariate Analysis*, **30**, 27-33. (**MR** #90m: 62073).

70. Babu, G. J. (1989). Applications of Edgeworth expansions to bootstrap – A review. *Statistical Data Analysis and Inference*, Y. Dodge (Ed.), Elsevier science publishers B. V., Amsterdam, 223-237. (**MR** #91m: 62076).

71. Babu, G. J., and Rao, C. R. (1988). Joint asymptotic distribution of marginal quantiles and quantile functions in samples from a multivariate population. *J. Multivariate Analysis*, **27**, 15-23. (**MR** #90e: 62078).

72. Babu, G. J. (1988). A note on comparison of conditional means. *J. Statistical Planning and Inference*, **19**, 253-259. (**MR** #89i: 62054).

73. Babu, G. J., and Bose, A. (1988). Bootstrap confidence intervals. *Statistics and Probability Letters*, **7**, 151-160. (**MR** #90c: 62047).

74. Patil, G. P., Babu, G. J. Hennemuth, R. C., Meyers, W. L., Rajarshi, M. B., and Taillie, C. (1988). Data-based sampling and model-based estimation for environmental resources. *Handbook of Statistics*, vol. 6, *Eds.*, P. R. Krishnaiah and C. R. Rao. Elsevier Science Publishers B. V. (North-Holland), Amsterdam, 489-513. (**MR** #90m: 62033).

75. Babu, G. J. (1986). Efficient estimation of the reciprocal of the density quantile function at a point. *Statistics and Probability Letters*, **4**,133-139. (**MR** #87e: 62041).

76. Babu, G. J. (1986). A note on bootstrapping the variance of the sample quantile. *Ann. Inst. Statist. Math.*, **38**, Part A, 439-443. (**MR** #88a: 62092).

77. Babu, G. J. (1986). Estimation of density quantile function. *Sankhyā*, Series A, **48**, 142-149. (**MR** #89h: 62057).

78. Babu, G. J., and Singh, K. (1985). Edgeworth expansions for sampling without replacement from finite populations. *J. Multivariate Analysis*, **17**, 261-278. (**MR** #87h: 62027).

79. Ghosh, M., Parr, W., Singh, K., and Babu, G. J. (1984). A note on bootstrapping the sample median. *Annals of Statistics*, **12**, 1130-1135. (**MR** #86e: 62027).

80. Babu, G. J., and Singh, K. (1984). On one term Edgeworth correction by Efron's Bootstrap. *Sankhyā*, Series A, **46**, 219-232. (**MR** #86g: 62053).

81. Babu, G. J., and Singh, K. (1984). Asymptotic representations related to jackknifing and bootstrapping L-statistics. *Sankhyā*, Series A, **46**, 195-206. (**MR** #86m: 62084).

82. Babu, G. J. (1984). Bootstrapping statistics with linear combinations of chi-squares as weak limit. *Sankhyā*, Series A, **46**, 85-93. (**MR** #87a: 62048).

83. Babu, G. J., and Singh, K. (1983). Inference on means using the bootstrap. *Annals of Statistics*, **11**, 999-1003. (**MR** #84i: 62049).

84. Babu, G. J. (1983). On the law of iterated logarithm for occupation measures of empirical processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, **65**, 73-81. (**MR** #85h: 60041).

85. Babu, G. J., and Singh, K. (1982). On r-quick limit sets for empirical and related processes based on mixing rv's. *J. Multivariate Analysis*, **12**, 508-525. (**MR** #84d: 60049).

86. Babu, G. J. (1982). Distribution of the values of $\omega$ in short intervals. *Acta Math. Acad. Sci. Hungar.*, **40**, 135-137. (**MR** #84e: 10062).

87. Deo, C. M., and Babu, G. J. (1981). Probabilities of moderate deviations in a Banach space. *Proceedings of the American Mathematical Society*, **83**, 392-397. (**MR** #83a: 60044).

88. Babu, G. J. (1981). On the mean-values and distributions of arithmetic functions. *Acta Arithmetica*, **40**, 63-77. (**MR** #84d: 10054).

89. Babu, G. J. (1980). On the distribution of multiplicative functions. *Acta Arithmetica*, **36**, 331-340. (**MR** #82a: 10060).

90. Babu, G. J. (1980). An inequality for moments of sums of truncated $\phi$ - mixing random variables. *Sankhyā*, Series A, **42**, 1-8. (**MR** #84g: 60028).

91. Babu, G. J., Ghosh, M., and Singh, K. (1978). On rates of convergence to normality for mixing processes. *Sankhyā*, Series A, **40**, 278-294. (**MR** #81m: 60034).

92. Babu, G. J., and Singh, K. (1978). Probabilities of moderate deviations for some stationary strong mixing processes. *Sankhyā*, Series A, **40**, 38-43. (**MR** #82k: 60056).

93. Babu, G. J., and Singh, K. (1978). On probabilities of moderate deviations for dependent processes. *Sankhyā*, Series A, **40**, 28-37. (**MR** #82i: 60047).

94. Babu, G. J., and Singh, K. (1978). On deviation between empirical and quantile processes for dependent random variables. *J. Multivariate Analysis*, **8**, 532-549. (**MR** #84g: 60054).

95. Ghosh, M., and Babu, G. J. (1977). Probabilities of moderate deviations for some stationary $\phi$ - mixing processes. *Annals of Probability*, **5**, 222-234. (**MR** vol. 55 #9235).

96. Babu, G. J. (1977). Some results on the distribution of additive arithmetic functions - I. *Sankhyā*, Series A, **39**, 1-10. (**MR** vol. 58 #16577).

97. Erdős, P., Babu, G. J., and Ramachandra, K. (1977). An asymptotic formula in additive number theory–II. *J. Indian Math. Soc.*, **41**, 281-291. (**MR** #81g: 10062).

98. Babu, G. J. (1976). Some results on the distribution of values of additive functions on the set of pairs of positive integers - II. *Acta Arithmetica*, **29**, 359-366. (**MR** vol. 55 #5564b).

99. Babu, G. J. (1976). Some results on the distribution of values of additive functions on the set of pairs of positive integers - I. *Acta Arithmetica*, **29**, 171-179. (**MR** vol. 55 #5564a).

100. Babu, G. J. (1976). On the distribution of arithmetic functions. *Acta Arithmetica*, **29**, 97-104. (**MR** vol. 53 #325).

101. Erdős, P., Babu, G. J., and Ramachandra, K. (1976). An asymptotic formula in additive number theory. *Acta Arithmetica*, **28**, 405-412. (**MR** vol. 53 #7974).

102. Babu, G. J., and Ghosh, M. (1976). A random functional central limit theorem for martingales. *Acta Math. Acad. Sci. Hungar.*, **27**, 301-306. (**MR** vol. 54 #11426).

103. Babu, G. J. (1975). Absolutely continuous distribution functions of additive functions $f(p) = (\log p)^{-\alpha}, \alpha > 0$. *Acta Arithmetica*, **26**, 401-403. (**MR** vol. 51 #10271).

104. Babu, G. J. (1975). Some results on the distribution of the values of multiplicative functions. *Sankhyā*, Series A, **37**, 386-395. (**MR** vol. 55 #7975).

105. Ghosh, M., Babu, G. J., and Mukhopadhyay, N. (1975). Almost sure convergence of sums of maximum and minimum of random variables. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, **33**, 49-54. (**MR** vol. 52 #1848).

106. Babu, G. J. (1974). On the characteristic function of the distribution of the values of additive arithmetic functions. *Lietuvos Matematikos Rinkinys*, **14**, 27-31. (**MR** vol. 52 #321).

107. Babu, G. J. (1973). Some results on the distribution of additive arithmetic functions–III. *Acta Arithmetica*, **25**, 39-49. (**MR** vol. 53 #324).

108. Babu, G. J. (1973). Some results on the distribution of additive arithmetic functions–II. *Acta Arithmetica*, **23**, 315-328. (**MR** vol. 53 #324).

109. Babu, G. J. (1973). A note on the invariance principle for additive functions. *Sankhyā*, Series A, **35**, 307-310. (**MR** vol. 58 #27877).

110. Babu, G. J. (1972). On the distribution of additive arithmetical functions of integral polynomials. *Sankhyā*, Series A, **34**, 323-334. (**MR** vol.50 #2109).

**Papers in conference proceedings and books:**

111. Martínez-Gómez, E. and Babu, G. J. (2010). A Relation Between the Exoplanets and Their Host Stars Based on Multivariate Analysis. In *2010 Joint Statistical Meetings Proceedings, Section on Physical and Engineering Sciences*, Alexandria, VA: American Statistical Association, 2387-2400.

112. Loredo, T. J.; Babu, G. J.; Borne, K.D.; Feigelson, E. D.; Gray, A. (2010). The Spectrum of LSST Data Analysis Challenges; Kiloscale to Petascale. Poster presented at 215th AAS meeting in 3-7 January 2010 in Washington, D.C. Presentation Number 401.24.

113. Martínez-Gómez, E. and Babu, G. J. (2009). A statistical model for the relation between exoplanets and their host stars. In the *Proceedings of the 24th International Workshop on Statistical Modelling*, Cornell University, James G. Booth (Ed.), 237-242.

114. Babu, G. J. and Feigelson, Eric D. (2006). Astrostatistics: Goodness-of-Fit and All That! In *Astronomical Data Analysis Software and Systems XV*, ASP Conference Series, C. Gabriel, C. Arviset, D. Ponz and E. Solano, (eds.), **351**, 127-136.

115. Jacob, J. C.; Katz, D. S.; Miller, C. D.; Walia, H.; Williams, R. D.; Djorgovski, S. G.; Graham, M. J.; Mahabal, A. A.; Babu, G. J.; vanden Berk, D. E.; Nichol, R. (2005). Grist: Grid-based Data Mining for Astronomy. In *Astronomical Data Analysis Software and Systems XIV*, ASP Conference Series, P. L. Shopbell, M. C. Britton, and R. Ebert (eds.), **347**, 306-310.

116. Graham, M. J.; Djorgovski, S. G.; Mahabal, A. A.; Williams, R. D.; Babu, G. J.; Feigelson, E. D.; vanden Berk, D. E.; Nichol, R.; Wasserman, L. (2005). VOStat: a distributed statistical toolkit for the Virtual Observatory. In *Astronomical Data Analysis Software and Systems XIV*, ASP Conference Series, P. L. Shopbell, M. C. Britton, and R. Ebert (eds.), **347**, 394-398.

117. Babu, G. J. (2004). Model fitting in the presence of nuisance parameters. In proceedings of *Astronomical Data Analysis-III*. Fionn D. Murtagh (Ed.). Electronic Workshops in Computing (eWiC). http://www.bcs.org/server.php?show=ConWebDoc.3901

118. Feigelson, E. D., and Babu, G. J. (2004). Statistical Challenges in Modern Astronomy. In *PhyStat2003: Statistical problems in Particle Physics, Astrophysics, and Cosmology*, 1-7, L. Lyons, R. Mount and R. Reitmeyer (Eds.), Stanford Linear Accelerator Center, Stanford, CA.

119. Mahabal, A.; Djorgovski, S. G.; Graham, M.; Williams, R.; Feigelson, E.; Babu, G. J.; Nichol, R.; Vanden Berk, D.; Wasserman, L. (2004). Doing Science with VOStat. *Bulletin of the American Astronomical Society*, **36**, p.806.

120. Babu, G. J., and McDermott, James P. (2002). Statistical methodology for massive datasets and model selection. In *Astronomical Data Analysis II*, Jean-Luc Starck and Fionn D. Murtagh (Eds.), Proceedings of SPIE, Vol. **4847**, 228-237.

121. McDermott, J. P.; Babu, G. J.; Feigelson, E. D. (2002). Low-Storage, Sequential, Simultaneous Estimation of Multiple Quantiles for Massive Datasets. *Bulletin of the American Astronomical Society*, **34**, p.742.

122. Mahabal, A. A.; McDermott, J. P.; Babu, G. J.; Feigelson, E. D.; Djorgovski, S. G.; Nichol, R.; Wasserman, L. (2002). Implementing Astrostatistics in the Virtual Observatory. *Bulletin of the American Astronomical Society*, **34**, p.743.

123. Babu, G. J., and Feigelson, E. D. (2001). Statistical methodology for National Virtual Observatory. In *Virtual Observatories of the Future*, 272-278, Robert J. Brunner, S. George Djorgovski, and Alexander Szalay (Eds.), Astronomical Society of the Pacific Conference Series Volume **225**, San Francisco.

124. Babu, G. J., and Manstavičius, E. (1999). Limit theorems for random permutations. In *Paul Erdős and his Mathematics (Research Communications)*, A. Sali, M. Simonovits, and V. T. Sós (Eds.), János Bolyai Mathemaical Society, Budapest, 19-22.

125. Feigelson, E. D., and Babu, G. J. (1997). Improving the statistical methodology in astronomy. In *Data Analysis in Astronomy V*, (V. di Gesú, M. J. B. Duff, A. Heck, M. C. Maccarone, L. Scarsi & H. U. Zimmermann eds.), World Scientific, Singapore, 281-288.

126. Babu, G. J., and Feigelson, E. D. (1997). The resurgence of astrostatistics. Proceedings of 51st session of the International Statistical Institute (Istanbul, 18-26 August 1997). *Bulletin of the International Statistical Institute*, Book 2, 371-374.

127. Feigelson, E. D., and Babu, G. J. (1997). Statistical methodology for large astronomical surveys. In *New Horizons from Multi-wavelength Sky Surveys*, (B. McLean et al., eds.) IAU Symp. 179, Kluwer, 363-370.

128. Feigelson, E. D., and Babu, G. J. (1992). Improving the statistical methodology of astronomical data analysis. *Astronomical data analysis software and systems I*, Diana M. Worrall, Chris Biemesderfer, and Jeannette Barnes (Eds.), Astronomical Society of the Pacific Conference Series, Vol **25**, 237-244.

129. Babu, G. J. (1992). Invited discussion of the paper "The Promise of Bayesian inference for astrophysics" by T. J. Loredo. In *Statistical Challenges in Modern Astronomy;* Eds. E. D. Feigelson and G. J. Babu. Springer-Verlag, New York, 298-299.

130. Babu, G. J., and Feigelson, E. D. (1992). CI in astronomy. In *Combining information: Statistical issues and opportunities for research.* Published by National Research Council. National Academy Press, Washington, D. C., 74-78.

131. Babu, G. J., Pennington, M., and Patil, G. P. (1986). Estimation of relative fishing power of different vessels. In Proceedings of the *Special Symposium on U.S. National Monitoring Strategies, Oceans 86.* Washington D.C. Vol. **3**, 914-917.

132. Patil, G. P., Babu, G. J., Boswell, M. T., Chatterjee, K., Linder, E., and Taillie, C. (1986). Statistical issues in combining ecological studies, with examples in marine fisheries research and management. Invited discussion paper presented at the ASA and EPA Conference on *Statistical Issues in Combining Environmental Studies*, at Washington D. C., 70-88.

**Book Reviews by G. J. Babu:**

1. Introduction to empirical processes and semiparametric inference, by M. R. Kosorok, Springer Series in Statistics. Springer-Verlag, New York, 2008. *Mathematical Reviews* (2012). **MR** #2012b: 62005.

2. Nonparametric goodness-of-fit testing under Gaussian models, by Yu. I. Ingster and I. A. Suslina, Lecture Notes in Statistics, 169. Springer-Verlag, New York, 2003. *Mathematical Reviews* (2005). **MR** #2005k: 62003.

3. R. R. Bahadur's lectures on the theory of estimation, edited and with a preface by Stephen M. Stigler, Wing Hung Wong and Daming Xu. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 39. Institute of Mathematical Statistics, Beachwood, OH, 2002. *Mathematical Reviews* (2005). **MR** #2005a: 62008.

4. Bootstrap Techniques for Signal Processing, by Abdelhak M. Zoubir and D. Robert Iskander, Cambridge University Press, Cambridge, UK., 2004. *Technometrics*, **47**, No. 3 (2005) 374–375.

5. Multivariate Permutation Tests, by Fortunato Pesarin, John Wiley, West Sussex, UK., 2001. *Technometrics*, **44**, No. 3 (2002) 290.

6. Fundamentals of Modern Statistical Methods, by Rand R. Wilcox, Springer-Verlag, New York, 2001. *Technometrics*, **44**, No. 1 (2002) 83.

7. Generalized Method of Moments Estimation, edited by Lászó Mátyás, Cambridge University Press, Cambridge, 1999. *Mathematical Reviews* (2000). **MR** #2000h: 62128.

8. Subsampling, by Dimitris N. Politis, Joseph P. Romano, Michael Wolf, Springer-Verlag, New York, 1999. *Sankhyā*, Series A, **62** (2000) 152–153.

9. Bootstrap Methods: A practitioner's guide, by Michael R. Chernick, Wiley-Interscience, 1999. *Mathematical Reviews*, (2000). **MR** #2000k: 62083.

10. Introduction to Probability Models, Sixth Edition, by Sheldon M. Ross, Academic Press, San Diego, 1997. *Technometrics*, **40** (1998) 78.

11. The Jackknife and Bootstrap, by J. Shao and D. Tu; Springer-Verlag, New York, 1995. *Mathematical Reviews* (1997). **MR** #97b: 62070.

12. Asymptotic Statistics, by R. N. Bhattacharya and Manfred Denker; Birkhäuser, Boston, 1990. *Mathematical Reviews* (1994). **MR** #94d: 62042.

13. The bootstrap and Edgeworth expansion, by Peter Hall; Springer-Verlag, New York, 1992. *Mathematical Reviews* (1993). **MR** #93h: 62029.

14. Asymptotic Theory of Statistical Inference, by B. L. S. Prakasa Rao; Wiley, New York, 1987. *Journal of the American Statistical Association*, **83**, No.404 (1988) 1217–1218.

15. Arithmetic Functions and Integer Products, by P. D. T. A. Elliott; Springer-Verlag, New York, 1985. *Mathematical Reviews* (1986). **MR** #86j: 11095.

16. A method for the derivation of limit theorems for some classes of dependent random variables, by Lothar Heinrich; Wissenschaftliche Sitzungen zur Stochastik [Scientific Meetings on Stochastics], 83-5. Akademie der Wissenschaften der DDR, Institut für Mathematik, Berlin, 1983. 41 pp. *Mathematical Reviews* (1986). **MR** #86b: 60040.

17. Advances in Probability Theory: Limit Theorems and Related Problems, Edited by A. A. Borovkov; Optimization Software Inc., New York, 1984. *Sankhyā*, Series A, **47** (1985) 419.

18. Martingale Limit Theory and its Application, by C. C. Heyde and P. Hall; Academic Press, New York, 1980. *Sankhyā*, Series A, **45** (1983) 397–398.

19. Probabilistic Number Theory - II: Central limit theorems, by P. D. T. A. Elliott; Springer-Verlag, New York, 1980. *J. Mathematical and Physical Sciences* (1982).

20. Probabilistic Number Theory - I: Mean-value theorems, by P. D. T. A. Elliott; Springer-Verlag, New York, 1979. *J. Mathematical and Physical Sciences* (1982).

21. Success Epochs in Bernoulli Trials (with applications in number theory), by W. Vervatt; Mathematisch Centrum, Amsterdam, 1972, *Sankhyā*, Series A, **35** (1973) 112–113.

**Selected recent research and future plans; highlights of earlier research, interdisciplinary work, honors, and professional & university service.**

**Gutti Jogesh Babu**

I have broad research interests in both statistics and probability, and in their applications to massive data, problems in biomedical research, astronomy and astrophysics. I have contributed extensively to probabilistic number theory (my **Erdős** Number is **1**; http://www.oakland.edu/enp/; https://files.oakland.edu/users/grossman/enp/Erdos1.html), resampling methods, nonparametric methods and asymptotic theory. I have published seven books and over 130 research articles in leading journals. An eighth book on Statistical Inference is under preparation.

In addition to theoretical research in statistics and probability, I continue to collaborate on Grid Data Mining for Astronomy project and on astrostatistical issues for Large Synoptic Survey Telescope. I have recently initiated collaboration with the Atacama Large Millimeter/Submillimeter Array (ALMA) group at the National Radio Astronomy Observatory (NRAO), on statistical methods for faint source detection/transient detection in multilayer astronomical datacubes. ALMA is expected to produces massive data cubes that require parallel computing. This work on astronomical data cubes lends itself for parallelization for computational purposes. The methods will also apply to the very important problem of detection of transients in multi-epoch survey data. The Keck Institute for Space Studies at Caltech has very recently drafted me as a core team member. The main purpose is to develop strategies (statistical methods and computationally efficient algorithms) that enhance the sensitivity for some types of computationally limited signals, where either faster computers or better algorithms would lead to more discoveries in the same massive data sets or data streams. This is especially important in view of the current push for analysis of 'BIG DATA' by the White House.

Organized and led a semester long research program on Astrostatistics at Statistical and Applied Mathematical Sciences Institute (SAMSI) in 2006. I will be again leading a second Astrostatistics program at SAMSI in 2012. I have organized and led very popular annual summer schools in statistics for astronomers at Penn State since 2005 and three in collaboration with the Indian Institute of Astrophysics. The Penn State Summer Schools have trained over 420 participants since its inauguration in 2005. I led the development of Web computing environment VOStat specifically for very-large datasets encountered in astronomy research in general, and Virtual Astronomical Observatory (VAO) in particular. VOStat is in the process of being integrated seamlessly with VO services for the use by the international VAO community.

I am Vice-Chair and an Executive Board Member of the recently constituted, International Statistical Institute's Astrostatistics Committee. The committee's main goal is to foster collaboration between members of the statistical and astronomical communities. I am one of the *five core team members* of the **LSST** *Informatics and Statistics Science Collaboration*, the only new collaboration approved in several years. I am the only statistician on the core team. The Office of NSF Director (OD) invited me to be on the working group on US/India Virtual Institutes for Computational and Data-Enabled science and engineering in Bangalore in Dec 21-22, 2011. I was on the Panel on Data-Intensive Computing and Astrophysics.

# Research

## Interdisciplinary research

My current interdisciplinary research efforts focus on statistical methodology for the National Virtual Observatory (VO), Grid Data Mining for Astronomy (GRIST) on faint source detection methodology, and on astrostatistical issues for Large Synoptic Survey Telescope (LSST), in particular on Weak Lensing related issues.

I led an interdisciplinary team of astronomers, computer scientists and statisticians from Penn State, California Institute of Technology, and Carnegie Mellon University, to develop statistical methodology for the VO. The effort was funded by **my** NSF grant of over $1 million from the Division of Mathematical Sciences (I was the first one in the Department of Statistics to bring in such a large grant). This effort covers several areas of statistics including the fields of multivariate analysis, nonparametrics, Bayesian analysis, spatial point processes, density estimation and data mining. Under this project many methods and tools including: multidimensional goodness-of-fit tests for comparison of multivariate astronomical datasets with astrophysical models; and a preliminary version of VOStat, a web based service providing a suite of tools allowing astronomers to use both simple and sophisticated statistical routines on large datasets were developed.

**LSST**: I am a member of the Weak Lensing Science Collaboration team of LSST and I am one of the *five core team members* of the LSST *Informatics and Statistics Science Collaboration*. The 8.4-meter LSST will survey the entire visible sky deeply in multiple colors every week with its three-billion pixel digital camera, probing the mysteries of Dark Matter and Dark Energy, and opening a movie-like window on objects that change or move. This color 'movie' of the Universe will open an entirely new window: the time domain. Data Data everywhere! LSST will generate more than 30 Terabytes per night leading to a 50 Petabyte archive over 5 years; approximately $10^9$ static and time-variable sources will be observed again every 3 nights. My future research plans include, scaling the recently developed faint object detection techniques for the LSST mission, in addition to developing other astrostatistical methodology for LSST both at the design stage and at the analysis stage.

**NRAO**: Recently initiated collaboration with National Radio Astronomy Observatory (NRAO). The fields of radio and microwave astronomy are experiencing huge advances in instrumentation. Innovations in broadband receiver, fast correlator and other technologies are leading to order-of-magnitude improvements in sensitivities and throughput from meter through submillimeter wavelengths. The NSF-funded National Radio Astronomy Observatory (NRAO) is now constructing the Extended Very Large Array (EVLA). According to one cosmological model, the universe is dominated at its largest scales by 'dark energy' that makes it expand faster, and on the scale of galaxies and clusters of galaxies, gravitational effects from 'dark matter'. One of the predictions of this model, which has not been verified by observation, is that there should be many sub-galactic clumps of dark matter, which should be visible as 'dwarf' galaxies due to the influence of gravity. Dwarf galaxies are numerous but not in the numbers predicted by the theory. Observations of the 21 cm emission line of atomic hydrogen in nearby galaxies are a powerful diagnostic tool for studying the dynamics and gas distribution in these galaxies. The data comes in the form of 3-dimensional 'datacube' (or 'hyperspectral image' ) giving brightness as a function of location in the two-dimensional sky at many channels of frequency. The new EVLA and ALMA correlators

will typically produce $\sim 1-100$ GBy datacubes with images of $\sim 1-10$ million pixels at $\sim 1-10$ thousand frequency channels. Over a decade, ALMA will produce petabytes of datacubes. My plans include the use these data and the data from forthcoming missions to search for hydrogen associated with the missing dark matter clumps; these image cubes will be the inputs to the statistical source finding techniques that I have been developing.

**GRIST**: I was Co-PI for the $3 million NSF grant 'ITR: Grid Service Workflow System as a Research Environment for Science with Massive Data Sets'. The project is for the development of a scientific research environment suited for the exploration of massive and complex datasets, that would exploit the power of the modern IT infrastructure. I am the only statistician of the interdisciplinary team that included astronomers and computer scientists from Caltech, and Jet Propulsion Laboratory. The scientific motivation of Grist derives from creation and mining of wide-area federated images, catalogs, and spectra. An astronomical image collection will generally cover an area of sky several times – in different wavebands, different times, etc – and the data analysis should combine these multiple observations to a unified (federated) understanding of the physical processes in the Universe. With large scale panchromatic synoptic surveys becoming more common, image co-addition has become even more necessary as new observations start to get compared with co-added fiducial sky in real time. I have developed a robust object detection technique capable of detecting faint sources, even those not seen at all epochs which will normally be smoothed out in traditional methods. As we refine the method for large data sets, we are starting to incorporate it in to the Palomar-QUEST real-time pipeline. The methodology will also set a standard for the forthcoming synoptic surveys like Pan-STARRS and LSST. The future work in this area include multiple testing such as false discovery rates on clustered data.

## Statistics and Probability research

**Goodness of fit**: Many nonparametric goodness-of-fit tests, such as Kolmogorov-Smirnov, and Cramer-von Mises, are based on the empirical distribution function. In the presence of nuisance parameters, the tests are generally constructed by first estimating the nuisance parameters. However, even when the parametric model is specified, the asymptotic null distribution of the test statistic depends in a complex way on the unknown parameters. Bootstrap methods are used to estimate the null distribution. I have demonstrated that, under very general conditions, the difference between the empirical process and the population distribution with estimated parameters converges weakly to the same Gaussian process as the corresponding bootstrap version. This result is used to show that the bootstrap method consistently estimates the null distributions of various goodness-of-fit tests. These results hold not only in the univariate case but also in the multivariate setting. For the case when the hypothesis is rejected, using Kullback-Leibler measure of separation, I have developed a resampling method to set confidence bands to the difference of the true and the closest distribution in the specified family. The ideas generated here would also help in model selection. I have developed a model selection procedure based on jackknife type arguments, as an alternative to *Akaike Information Criterion* (AIC) and *Bayesian Information Criterion* (BIC), These results are very useful to problems in astronomy.

**Multivariate density estimation**: A multivariate version of Bernstein polynomials for approximating a bounded and continuous function is adapted for smooth estimation of a distribution function concentrated on the unit hypercube. The smoothness of the resulting estimator, naturally

lends itself in a smooth estimator of the corresponding density. The functions with other compact or non-compact support are also dealt with through suitable transformations. The asymptotic properties, namely, strong consistency and asymptotic normality of the resulting estimators are investigated under strong mixing. The problem was motivated by estimation of conditional densities in non-linear dynamical systems. Metric entropy is an important measure of chaos in a dynamical system. A new statistical method is developed for *(Chaotic processes)* filtering and entropy estimation of a chaotic map from noisy data.

**Applied Probability**: The Ewens sampling formula gives the distribution of the allelic partition of a sample of genes from the so-called infinitely many neutral alleles model of population genetics. Though for the past several years the focus has moved away from allozyme data and more towards DNA sequence data, the combinatorial content of the Ewens sampling formula has been recognized as central to the study of a broad class of combinatorial structures in applied probability, and it is closely associated with partitions of an integer. Continuing my work on functional limit theory for 'certain' processes defined through partial sums of dependent variables with respect to the Ewens sampling formula, I have shown that the process converges weakly in a function space if and only if a related process defined through sums of independent random elements converges. Recently, using techniques from probabilistic number theory, I have developed a functional limit theory where the limiting processes need not be processes with independent increments. For these processes, under Ewens sampling formula, the limiting process of the partial sums of dependent variables differs from that of the associated process defined through the partial sums of independent random variables. There are many interesting problems in this area. The work will continue. Although the results are motivated by Ewens sampling formula, they have wide range of applications including partitions of an integer in number theory and other combinatorial structures such as, 'assemblies'. These objects are related to physics (representations and Young diagrams), theoretical computer science (tree-based searching and symbolic processing algorithms based upon forests), cryptology (factorization of polynomials), and chemistry (random trees as models for cyclic polymerization).

**Bootstrapping empirical measures**: Contrary to the intuition, properly normalized empirical distribution is not approximated uniformly by its bootstrapped counterpart. The bootstrap does not work for variables given by rare events or very small quantiles with Poisson limit laws. The main problem is at the values near the tails of the population distribution. My results specify the central regions where the uniform distance vanishes if the sample size tends to infinity.

**Asymptotics for functions of marginal quantiles:** While I collaborated with C. R. Rao on the large-sample properties of marginal sample quantiles in the case of independent vectors, I along with Bai and Choi of National University of Singapore, investigated the large-sample properties of means of functions of marginal quantiles. The investigation is motivated by the problem of parameter estimation in regression models when the linkage between the dependent and independent variables is partially lost. The study includes asymptotic normality, the strong law of large numbers, and functional limit theory.

**Other research includes**: a) Robust tests for unrestricted and ordered alternatives in the multi-sample problem without assuming homogeneity of scales and/or symmetry of the underlying distributions. The methodology consists of bootstrapping appropriately centered Mann-Whitney statistics. Data sets from Physics and Psychology illustrate the methodology. b) Simultaneous estimation of multiple quantiles for massive streaming datasets with applications to density estimation. c) Estimation of occurrence/exposure rates under competing risks.

## Interdisciplinary Promulgation

**Center for Astrostatistics**: Astronomy at the beginning of the 21st century, and particularly research arising from wide-field survey observatories at various wavebands, finds itself with serious challenges in statistical treatments of data to achieve its astrophysical goals. A vast range of statistical problems arise in the scientific interpretation of astronomical studies. It is this diversity of statistical issues confronting astronomy today that led to the creation of the Center for Astrostatistics in 2003 at Penn State to facilitate development and promulgation of statistical expertise and toolkits for astronomy and related observational sciences. I am the founding director of the Center for astrostatistics. The Center serves as a crossroads where researchers at the interfaces between statistics, data analysis, astronomy, space and observational physics collaborate, develop and share methodologies, and together prepare the next generation of researchers. I also maintain the resource rich Center Web site that includes: bibliographies, summer school lectures, tutorials, data sets, StatCodes, and VOStat. Due to the multifaceted activities, the Penn State is now known as THE place for astrostatistics world wide. Due mainly of the efforts of the center for astrostatistics, several national as well as international collaborating groups in the field have been established around the world. Astronomers and astrophysicists around the globe tap on the resources of the Center for astrostatistics for guidance on Statistical methodology.

**SAMSI**: At the invitation of Statistical and Applied Mathematical Sciences Institute (SAMSI), I directed a semester-long Astrostatistics program in Spring 2006. The principal purpose of the SAMSI program on Astrostatistics is to identify promising research paths for statistical sciences and applied mathematics in problems of observational astronomy, astrophysics and particle physics, and to initiate research on these problems. A vital ingredient of the program was to provide a single geographical location, where researchers at the interface between statistics, applied mathematics, astronomy, and particle physics can congregate and initiate lasting collaborations. The participation by graduate students and postdocs gave them a rare opportunity to develop skills needed for cross-disciplinary work. I organized several research working groups (Exoplanets, Surveys, Gravitational lensing, Source Detection & Feature Detection, and Particle Physics). They met throughout the semester at SAMSI. The program was concluded with the interdisciplinary conference 'Statistical Challenges in Modern astronomy IV' at Penn State. Many technical reports emerged. The program helped spawn many new collaborative teams around the globe. I will be again leading a second Astrostatistics program at SAMSI in 2012 as part of '*2012-13 Program on Statistical and Computational Methodology for Massive Datasets*'.

**Astrostatistics Schools**: Under my direction, the Center for Astrostatistics, organized very popular annual summer schools since 2005 at Penn State to train astronomers and physicists in advanced statistical methods for handling a diversity of statistical issues confronting astronomy, space sciences, and high energy particle physics. I have also organized, jointly with the Indian Institute of Astrophysics, a 6-day school in statistical inference for practicing astrophysicists and physicists affiliated with Indian institutions in India in July 2007, 2008 and July 2010. At the invitation of the Instruments Division of the Space Telescope Science Institute (STScI), I presented series of lectures on probability and statistics in Fall 2011 at STScI.

**Statistical Challenges in Modern Astronomy (SCMA)**: The SCMA conferences, held every five years since 1991, are the premiere forum for research statisticians and astronomers to discuss methodological issues of mutual interest. President of the American Statistical Association, Bradley

Efron, in the 'President's Corner' of the Jan 2004 issue of Amstat News refers to the success and the impact of the Penn State's 'Astrostatistics' conference series:

> "The statistical tide continues to roll in, now lapping at the previously unreachable shores of the hard sciences. This September I attended 'Phystat2003' at the Stanford Linear Accelerator Center, a joint conference for physicists and statisticians. Phystat2003 was at least partially inspired by the success of Penn State's 'Astrostat' conference series. Yes, confidence intervals apply as well to neutrino masses as to disease rates, and raise the same interpretative questions, too."

Continuing this tradition, the fifth in a series of interdisciplinary international research conferences, SCMA V was organized in June 2011. It attracted leading statisticians and astronomers to Penn State.

## Honors and Awards:

I was elected Fellow of the Institute of Mathematical Statistics in 1987 for his work on the asymptotic theory of bootstrap methods, and was elected Member of the International Statistical Institute in 1989. I was elected Fellow of American Statistical Association in 1997, for outstanding and wide-ranging contributions to probability and statistics; for leadership in promoting interdisciplinary activities to bring astronomers and statisticians together; and for service to the statistical profession. I was elected Fellow of American Association for the Advancement of Science in 1997, for research on asymptotic theory, resampling methods, probabilistic number theory, and statistical methods for astronomy and for promoting interdisciplinary activities. I have also received National Research Council's Twinning fellowship for 1997-1999 to initiate collaboration, on Statistical Group Theory and Probabilistic Number Theory, with colleagues from Vilnius University, Lithuania. I was appointed as Research Professor (Spring 2005), Mathematical Sciences Research Institute, Berkeley; and SAMSI University Fellow (Spring 2006). I received a 2007 University Continuing Education Association (Mid-Atlantic Region) award for exemplary non-credit program development for the 2007 summer school.

## Press and electronic media recognition

1. The interdisciplinary work received wide recognition in the press. Span (Celestial Collaborators, p. 56), New Delhi 2005 on U.S.-India cooperation in astronomy and astrophysics. SPAN is a publication brought out by the American Embassy in India.

2. Penn State Outreach Magazine, Spring 2007: Rock Stars, Real Stars & Astronomical Success.

3. Media accounts of research and outreach: Featured in the DVD produced in 2005 by Penn State Public Broadcasting entitled, 'Penn State Outreach. Shared Stories: Faculty talk about Outreach'

## Highlights of Professional and University Service

*Service to the profession*:

- Member of the IMS Committee on Fellows (2001-2004), and served as the chair of the committee in 2003.
- Chair, Program Leaders Committee, SAMSI Astrostatistics Program January - May 2006.
- Executive Board Member and Vice-Chair, International Statistical Institutes's Astrostatistics Committee. `http://isi.cbs.nl/COMM/AstroStat/index.htm`
- Member of the NSF Mathematical Sciences Postdoctoral Research Fellowships panel 2005 - 2006, and a member of the Executive committee for NSF MSP Postdoc panel 2006.
- Served on many NSF and NASA panels.
- Written over 200 reviews for 'Mathematical Reviews' and 'Zentralblatt'.

*Editorial responsibilities*: I am the founding Editor-in-Chief of the journal 'Statistical Methodology'. In addition, I was a coordinating Editor for 'Journal of Statistical Planning and Inference' till 2004, an associate editor for 'Journal of Nonparametric Statistics' and Co-Editor 'Sankhya' until December 2007.

*Master of Applied Statistics Program*: I have actively participated in developing the proposal for a Professional Masters degree program in the Penn State's Department of Statistics and funded by the Alfred P. Sloan Foundation. The grant form the Sloan Foundation has allowed the Department of Statistics to create and offer a professional degree program 'Master of Applied Statistics' (MAS). I have led the MAS program through the approval process of Penn State's Graduate School, and served as its **first Professor-in-charge** until 2007.

*College and University wide activities*: I chaired the Committee on Academic Standards of the Graduate Council during 2004-2005, and served on the Executive committee of the Graduate Council during 2003-2005. I served on the World Campus steering committee for several years until 2005, after being appointed to it by the Penn State's Executive VP & Provost. I also served on the Penn State Review Panel on Outreach Scholarship Conference (a partnership between Penn State, the University of Wisconsin-Extension, Ohio State, and the University of Georgia, 2006). I served as the Chair of the selection committee for the John M. Chemerda Lectures in Science (2003-2004) and was responsible for bringing John Nash to Penn State. I served on the Eberly College of Science Strategic Vision Committee (2004-2005), and continue to serve on the Eberly College of Science Outreach Council.

## Earlier research

### Probabilistic number theory

I contributed extensively to *Probabilistic number theory* in early 1970's, and published a monograph, **Probabilistic Methods in the Theory of Arithmetic Functions**, in 1978. My main

contributions in this area include a partial solution to a long standing conjecture of Erdős, and the result that every bounded additive arithmetic function has a singular distribution. I introduced a concept of density of natural numbers, capable of detecting large gaps in a set of natural numbers, and investigated the existence of distribution of values of an arithmetic function under this density. An important consequence of this is that if $\omega(m)$ denotes the number of distinct prime factors of $m$, then

$$\#\{n < m < n + b(n) : \omega(m) - \log\log m < x\sqrt{\log\log n}\}/b(n) \to \Phi(x), \qquad (*)$$

where $(\log\ b(n))\sqrt{\log\log n}/\log n \to \infty$, and $\Phi$ denotes the standard normal distribution function. This is a generalization of the well known Erdös-Kac Theorem and it leads to a better understanding of integers with large number of prime factors. It is shown that (*) fails to hold if $b(n) < (\log n)/(\log\log n)^2$. Given an additive function $f$, the problem of determining the slowest growing function $b$ so that $f$ has a distribution was also studied. Part of this was joint work with late Professor Paul Erdős.

## Bootstrap methods

My research in early 1980's provided theoretical support to Efron's resampling method called 'bootstrap'. This work on the asymptotic theory for the bootstrap method resulted in establishing the superiority of the bootstrap approximation for a wide class of statistics. I have shown that the bootstrap works best for asymptotically pivotal quantities (studentized versions). This is achieved using Edgeworth expansions. This laid the foundation for subsequent work on second order approximations of the bootstrap method. I was the **first** to show that bootstrap fails if the second moment is not finite.

I along with C. R. Rao and Pathak have proposed a sequential approach called Poisson bootstrap, in which resampling is carried out until a certain proportion of distinct observations are sampled from the original data. Using conditional Edgeworth expansions, we have established the second order correctness (skewness correction) for a wide class of statistics as in the case of classical bootstrap. One of the main advantages of the sequential approach over the fixed sample size bootstrap is, to prevent too many repeated observations in a bootstrap sample that may lead to a degenerate value for the statistic under consideration. Thus Poisson bootstrap avoids zero value for variance estimator.

Practically there is no literature on statistics which are asymptotically distributed as linear combinations of Chi-squares exists. To study such statistics I proposed a modification of bootstrap statistic and have shown that their distributions are very close to each other. This gave a much needed practical method to obtain confidence intervals for such statistics. This method was applied by various authors to study the so called U-statistics.

Bootstrap estimation of variance of sample quantiles were studied and exploited the proof to obtain a method of estimation of density quantile function. Simulation studies showed that the bootstrap method gives better results for studentized statistics. These observations are explained by the theoretical studies. A method to obtain confidence intervals is suggested using the bootstrap method. It is shown to perform better than the percentile method of Efron in the one sided case. The method is extended to autoregressive models.

*Subsample and half sample* methods are closely related to bootstrap and jackknife method. I investigated the large sample performance of this method. It is shown that the half-sample method is robust in estimating the parameters of a linear regression model when the errors are heterogeneous.

To assess the accuracy of the estimators thus obtained, one needs to estimate the variance of the estimators. There are several options, including bootstrap and halfsample methods. Bootstrap and jackknife are two widely used methods to estimate variance of a statistic. For non-smooth statistics such as sample median, it is well known that jackknife method fails. In this connection, I examined the bootstrap method using the notion of breakdown point in robustness, in the context of estimation of the variance of an estimator and of confidence intervals. Even when the estimators are robust, bootstrap estimator of the variance is strongly influenced by a single outlier, as the bootstrap utilizes all the data points. On the other hand halfsample method has high breakdown point, sometimes as high as 1/4, in the case of estimator of the variance of sample median.

## Edgeworth expansions

Edgeworth expansions play an important role in resampling methods such as bootstrap. In 1991, I obtained an s-term Edgeworth expansions for a wide class of statistics which are smooth functions of lattice and nonlattice marginals, where $s > 2$. The result is then applied to a statistic similar to the student's t-statistic, where the scaling factor, the sample standard deviation is replaced by the more robust mean absolute deviation. Edgeworth expansions for the product limit estimator and estimators based on product limit estimator were also obtained in the same year.

I and Z. D. Bai have established Edgeworth expansions for sums of independent but not identically distributed multivariate random vectors. We have also established expansions for functions of multivariate means under partial Cramér's condition and under minimal moment conditions. These results are applied to obtain valid Edgeworth expansions for estimates of regression parameters in linear errors-in-variable models. Using these expansions, the bootstrap distribution is shown to approximate the sampling distribution of the studentized estimators, better than the classical normal approximation. This justifies the use of bootstrap in applying the errors-in-variables regression to the cosmic distance scale, one of the important problems in Astronomy. Expansions that are local in one coordinate and global in rest of the coordinates were also obtained for sums of independent but not identically distributed random vectors. The results were then applied to derive Edgeworth expansions for bootstrap distribution, Bayesian bootstrap distribution, and for the distributions of statistics based on samples from finite populations. I explored expansions for sums of non-identically distributed random vectors and of random vectors with lattice and non-lattice coordinates, along with their applications to errors-in-variables models, least absolute deviation estimators etc.

*Edgeworth expansions for samples from finite populations:* Edgeworth expansions were obtained for the mean of a simple random sample drawn, from a finite multivariate population, without replacement. These are obtained under very general assumptions, which are easy to verify in practice. Consider the two sample non-parametric statistics of the type $T = (1/n)\sum f(R_i/N)$, where $f$ is a smooth function and $R_1, \ldots, R_n$ are the ranks of one of the samples. In our results we need only to assume that $f$ is continuous and monotone in a small interval. It appears that not much is known about the expansions of $T$, when $\{f(i/N)\}$ takes only two values. This and other

related problems in sample surveys can be handled by using the results on the lattice case. As one of the applications we obtain expansions for the univariate statistics which can be expressed in a certain linear plus a quadratic form. A fairly large class of statistics used in sample surveys fall in this category. These results can be used to get bootstrap approximation to various statistics in the finite population case.

## Other research

*Density quantile estimation:* Density quantile estimators based on the smoothness properties of the density were constructed and were shown to be asymptotically efficient in the mean square error sense. Unlike density estimators these do not require knowledge of the actual values of the derivatives of the density. Uniformly almost sure bounds for these estimators in an interval were obtained in the dependent case too.

*Competing risk models:* In a series of papers, a standard competing risk situation with possibility of censoring due to withdrawal or end of study were investigated. Large sample theory is derived and bootstrapping is discussed as a way to estimate the variance. Some of this is joint work with C. R. Rao.

*Moderate deviations in Banach spaces:* Probabilities of moderate deviations for i.i.d. sequences taking values in a separable Banach Space under precise necessary and sufficient conditions were obtained jointly with C. M. Deo. These results are not known earlier even for real valued random variables.

*Occupation measure of empirical processes:* The limit behavior of the occupation distribution

$$L_t(E) = \int_0^1 I_E(\sqrt{(\log\log t)/t}\ \ K(s,t))ds,$$

where $K$ denotes a Kiefer process, was studied. Strong approximation results are then used to derive the law of iterated logarithms in Chung's form for various functions of empirical processes.

*Robust estimation:* The large sample properties of statistics based on the robust mean absolute deviation from the sample mean as well as sample median are obtained. These are applied to regression context, the procedures are more robust against outliers compared to the usual procedures base on the classical least squares method.

*Mixing sequences:* The so called **r**-quick limit points of empirical distribution functions of mixing processes were characterized. Also obtained an **r**-quick version of Bahadur-Kiefer type representation for sample quantiles. These results are applied to linear functions of order statistics.

*Environmental statistics:* Trawl surveys are carried out regularly by the Northeast Fisheries Center (NEFC) to assess the fish stocks of various species. The external factors influencing the assessment include the ship used, doors, nets, etc. Occasionally some of these have to be replaced. So there is a need for a conversion factor to neutralize this influence. Major difficulty encountered in this problem is due to the large proportion of zero catches. A method of estimation for the conversion factor is suggested and various asymptotic properties of the estimator studied. The results are illustrated by a data set provided by NEFC.