

## Reports of IMS Standing Committees 2006

### Committee for the Carver Medal

	<u>Term expires</u>
Iain Johnstone (Chair)	Aug-06
David Siegmund	Aug-07
Stephen Stigler	Aug-08

This year's Carver Committee is delighted to award the 2006 Carver Medal to Robert V. Hogg, Professor Emeritus, Department of Statistics and Actuarial Science at the University of Iowa. A draft citation reads: "To Bob Hogg, whose past service as IMS Program Secretary (1969-1974) and long term support of the Central Region helped build and sustain the IMS and continues to epitomize the best in service to the Institute, we award the Carver Medal."

### Committee on Electronic Issues

	<u>Term expires</u>
Jim Pitman (Chair)	Aug-06
Arnold Stromberg	Aug-06
David Nualart	Aug-07
Qi-Man Shao	Aug-07
Vacant	
Vacant	
Bernard Silverman (ex officio)	Dec-07
Chris Burdzy (ex officio)	Dec-08
Paul Shaman (ex officio)	Aug-06

No report submitted.

### Committee on Fellows

	<u>Term expires</u>
Christian Robert (Chair)	Aug-06
Dorota Dabrowska	Aug-06
Brenda MacGibbon	Aug-06
Jianqing Fan	Aug-07
Jim Fill	Aug-07
John Marden	Aug-07
Uwe Einmahl	Aug-08
Daniel Pena	Aug-08
Sue Wilson	Aug-08

The IMS Committee on Fellows recommended the following members for election to Fellow. IMS Council approved the election in May 2006.

Naray Balakrishnan  
Peter Bühlmann  
Andreas Buja  
Tony Cai  
Alicia Carriquiry

Kung-Sik Chan  
Rong Chen  
Marie Davidian  
Ronald Doney  
Chris Genovese  
Subashis Ghosal  
Irene Gijbels  
James Hobert  
Susan Holmes  
Wai-Keung Li  
Wenbo Li  
David Madigan  
Naisyin Wang  
Martin Wells  
Heping Zhang

### **Committee on Finance**

	<u>Term expires</u>
Louis Chen (Chair) (ex officio)	Aug-06
David Scott	Aug-08
Hans Künsch	Aug-06
Mike Steele	Aug-07
Maury Bramson	Aug-08
Alan Karr	Aug-08
Per Mykland	Aug-08

No report submitted.

### **Committee on Memorials**

	<u>Term expires</u>
Ingram Olkin (Chair)	Aug-07
Alan Karr	Aug-06
Vacant	Aug-08
Bernard Silverman (ex officio)	Dec-07
Ed George (ex-officio)	Dec-07

No report submitted.

### **Committee on New Researchers**

	<u>Term expires</u>
Peter Craigmile (Chair)	Aug-07
Imola K. Fodor	Aug-06
Herbert Lee	Aug-06
Jeff Picka	Aug-06
Peter Song	Aug-06
Galin Jones	Aug-07
Xiaoming Sheng	Aug-08
Peter Hoff	Aug-08
Rebecka Jornsten	Aug-08

Committee on New Researchers  
Report of the Chair, Peter Craigmile  
(as of June 21, 2006)

Ninth New Researchers' Conference (2006 NRC)  
Program Chair: Peter Craigmile (The Ohio State University)  
Local Chair: Peter Hoff (University of Washington)

The New Researchers' Conference (NRC) is an annual IMS conference dedicated to creating interaction between new researchers in the fields of Probability and Statistics. The Ninth NRC will be held at the University of Washington in Seattle, August 1-5, 2006, immediately prior to the Joint Statistical Meetings, also to be held in Seattle. The website for the conference is at <http://www.stat.ohio-state.edu/~pfc/NRC/>. All participants are expected to give a short, expository talk or contribute a poster on their research.

In addition, the conference will consist of panel sessions of journal editors and funding agencies to discuss topics highly relevant to new researchers. Out of the people who applied 20 people, were turned away for the following reasons: 1 person who had attended the conference before; 7 people who will receive their PhD in 2007 or later; 12 people who missed the application deadline by at least a month.

This left "eligible" 92 participants that we invited to attend the meeting; 86 accepted. The attendees are from all over the world. Conference abstracts are available for viewing on the conference website. We are expecting 4 committee members to attend and help run the meeting.

#### Invited speakers and panels

The IMS Tweedie New Researcher Invited Lecture will be Savas Dayanik (Princeton University), who was selected by the IMS Travel Awards Committee. Savas receives funding to travel to both the 2006 NRC and JSM. Our confirmed invited speakers are Roger Berger (Arizona State University), Dalene Stangl (Duke University), and Bruce Weir (University of Washington). The journals panel will include Stephen Portnoy (Co-editor, JASA Theory and Methods), Jianqing Fan (Co-editor, The Annals of Statistics), Edward George (Co-editor, Statistical Science) and Naisyin Wang (Co-editor, Biometrics). The funding panel will include Michelle Wagner (National Security Agency) National Science Foundation, Mathematical Sciences - Statistics Patrick Heagerty (University of Washington, speaking about NIH funding).

#### Funding for the conference

We currently have funding from the NSF for \$20,000, from the NSA for \$10,000 and from the ONR for \$10,000. We did not receive the \$20,000 asked for from the NIH. The review reports were not bad, but there seems to be a shift to the funding of more focused NIH-related conferences. With help from Dr. Ram Tawari from the National Cancer Institute, we are applying for "Exception" support of \$5,000 from NIH. We will be informed about whether this is successful at the end of July.

Given the cut in NIH support, we will have to reduce the amount of travel support that we can provide to participants significantly. Committee members Peter Craigmile and Peter Hoff were instrumental in securing these funds. We received much needed support from Mayetri Gupta, Rebecka Jornsten, and Xiaoming Sheng, and administrative support from both the University of Washington and The Ohio State University.

#### The proposed 2007 New Researchers' Conference in Utah

Location: The University of Utah

Local Chair: Xiaoming Sheng

Funding: Funding will be requested from the same organizations that have funded previous NRCs.

The proposal to NSA for 2007 was submitted in October 2005.

## Committee on Nominations

	<u>Term expires</u>
Xuming He	Aug-06
Alessandra Rosalba Brazzale	Aug-06
Guido del Pino	Aug-06
Peter Glynn	Aug-06
Larry Goldstein	Aug-06
Andreas Greven	Aug-06
Geoffrey Grimmett	Aug-06
Chao Hsiung	Aug-06
Lynn Kuo	Aug-06
Makato Maejima	Aug-06
Thomas Mikosch	Aug-06
Rahul Mukerjee	Aug-06
Julia Norton	Aug-06
Lara Schmidt	Aug-06
Ruth Williams	Aug-06
Regina Liu (ex-officio)	Aug-06

### Report from the IMS Committee on Nomination

The IMS Committee on Nomination this year consisted of 15 members, chaired by Xuming He from the University of Illinois at Urbana-Champaign. The committee started working as soon as it was formed. By early 2006, it completed nomination of 1 candidate for the IMS President-Elect and 10 candidates for the IMS Council.

The first stage of the committee work involved discussion on specific qualifications that we would be looking for when identifying candidates. The committee placed emphasis on three factors when screening candidates: global perspective, stature in our research community as well as in the broader mathematics and science community, and active involvement with the IMS. When looking for new Council members, the committee kept diversity in mind. The final list of the candidates included people from North America, Europe, as well as Asia.

The committee was open to consider all the candidates suggested by the committee members as well as those suggested by other IMS members. However, only one candidate was formally suggested to the committee by IMS members not in the committee.

After initial discussions, the committee considered 4 candidates for the President-Elect. Each committee member ranked the candidates, and by a pre-determined voting scheme, the committee selected Professor Jianqing Fan as the final candidate. The committee considered a total of 21 possible candidates for the Council through email discussions. Then each committee member ranked those candidates, and by a predetermined voting scheme, the committee agreed on 14 top-ranked candidates. The chair of the committee contacted the candidates to ask for their willingness in the order of the committee ranking. Two candidates declined to be nominated, so at the end, the 12 top-ranked candidates were contacted, and 10 agreed to be nominated.

## Committee on Publications

	<u>Term expires</u>
Ruth Williams	Aug-06
Yi-Ching Yao	Aug-07
John Einmahl	Aug-08
Richard Vitale (ex officio)	Dec-06

Morris L. Eaton (ex-officio)	Dec-06
Jianqing Fan (ex-officio)	Dec-06
Michael Phelan (ex officio)	Dec-06
Bernard Silverman (ex officio)	Dec-07
Ed George (ex-officio)	Dec-07
Jiayang Sun (ex officio)	Aug-07
Paul Shaman (ex officio)	Dec-07
Greg Lawler (ex-officio)	Dec-08
Ed Waymire (ex-officio)	Dec-08

No report submitted.

### **Committee to Select Administrative Officers**

	<u>Term expires</u>
David Scott (Chair)	Aug-06
Thomas Selke	Aug-07
Liu, Jun	Aug-08

No report submitted.

### **Committee to Select Editors**

	<u>Term expires</u>
Bin Yu	Aug-06
Soren Asmussen	Aug-07
Steve Lalley	Aug-08
Wing Wong (LNMS Ad hoc)	Aug-06
Claudia Klueppelberg (LNMS Ad hoc)	Aug-06
Diane Lambert (AOS Ad hoc)	Aug-06
Hans Künsch (AOS Ad hoc)	Aug-06
Philip Dawid (AOAS Ad hoc)	Aug-06
David Madigan (AOAS Ad hoc)	Aug-06

No report submitted.

### **Committee on Special Lectures**

	<u>Term expires</u>
Maria Eulalia Vares (Chair)	Aug-06
Rosemary A. Bailey	Aug-06
Irene Gijbels	Aug-06
Probal Chaudhuri	Aug-07
Xiao-Li Meng	Aug-07
Ed Waymire	Aug-07
Regina Liu	Aug-08
Steffen Lauritzen	Aug-08
Jean-Francois Le Gall	Aug-08
Morris L. Eaton (ex-officio)	Dec-06
Jianqing Fan (ex-officio)	Dec-06
Ed George (ex-officio)	Dec-07
Ed Waymire (ex-officio)	Dec-08

Note: This report was submitted to and approved by IMS Council February 2006

-----  
The Committee on Special Lectures has been working since last November and we have now concluded our task involving the selection of:

- \* One Wald Lecturer (for 2007)
- \* One Rietz Lecturer (for 2007)
- \* Eight Medallion Lecturers (for 2008)

According to new rules the committee should also propose an assignment to sessions for each ML, with some flexibility. Here is the conclusion of our work and all indications from the SL Committee, that I now submit to you and to the IMS Council:

2007 WALD Lecture: Jim Berger (Duke)

-----  
2007 RIETZ Lecture: David Siegmund (Stanford)

-----  
2008 Medallion Lecturers (alphabetical order within categories):

Probability:

Martin Barlow (British Columbia),  
Zhi-Ming Ma (Inst. Applied Math, CAS, Beijing),  
Wendelin Werner (Paris-Sud)

Statistics:

George Casella (Florida),  
Mark Low (Pennsylvania)

Interdisciplinary:

Peter Bartlett (Berkeley),  
Mary Sara McPeck (Chicago),  
Alistair Sinclair (Berkeley)

Sessions assignments. Taking into account adequacy between typical audience and research areas of the speakers, we arrived to the following suggestion:

2008 IMS Annual Meeting (Joint with BWC08):

July 20-26, Singapore

Martin Barlow(P), Zhi-Ming Ma (P), Wendelin Werner (P), M. Low (S)

-----  
JSM 2008: 2 sessions

August 3-7, Denver, Colorado

G. Casella (S) , A. Sinclair (I)

-----  
ENAR 2008: 1 session

March 16-19, Arlington, VA

Mary Sara McPeck (I)

-----  
WNAR 2008: 1 session

Peter Bartlett (I)

-----  
Remark: The above proposal has been considered the one with the best fit for the available sessions (4-2-1-1). In terms of flexibility: interchanging Low and Sinclair, if needed, should work as well. Though this was not a general feeling, there is some concern if the lecture by Peter Bartlett (his main research topics is machine learning) would match properly to WNAR. The alternative would involve shifting P. Bartlett to IMS/BWC. This depends on other variables beyond committee's responsibility.

Some details. The committee work involved two steps:

1. Discussion and approval of a procedure
2. Search, discussion, selection in two rounds (for all lectures), assignment to sessions.

For the MLs, at the first round we selected 2 lecturers in Probability, 2 lecturers in Statistics and 1 in the Interdisciplinary category, guaranteeing the minimal requirements according to "Guidelines". The second round involved selection across categories. Though the outcome does not really show, during the discussion we had in mind the need to balance other issues, including gender. The initial list was elaborated with this in mind. The description of the available sessions was also provided to the members before the selection procedure.

We started the selection with a list of names containing:

Wald: 5 names (one woman)

Rietz: 6 names (two women)

ML

Probability: 15 names (two women)

Statistics: 16 names (one woman)

Interdisciplinary: 17 names (three women)

### **Committee on Travel Awards**

Term expires

Xuming He (Chair) Aug-06

Marina Vannucci Aug-08

Victor Perez-Abreu Aug-07

No report submitted.

Annals of Applied Probability  
Annual Report 2006  
June 23, 2006

In the nearly six months since assuming the editorship (January 1, 2006), there have been 112 submissions to AAP via the Electronic Journal Management System (EJMS). These may be classified as 4 Accepted and 5 Awaiting Revision versus 38 Rejected and 3 Withdrawn. The remaining 62 are Under Review. Based on this approximate submission rate (224 submissions/yr) and ignoring papers under review, the Acceptance/Rejection proportions suggests a slightly lower acceptance rate in comparison to the 65/142 split (32% accepted) for 2005. However the comparison is obviously tenuous at this stage. The 9 year historic submission rates are: 134, 183, 145, 170, 137, 149, 164, 196, 207. The previous editor Robert Adler noticed that there is evidence in the historic data on published papers for a possible conservation law of AAP papers; in that the acceptance rate has naturally varied in proportion to the number of submissions. There has been no explicit effort to control the number of publishable papers in AAP, beyond quality control standards imposed by referees and the editorial board.

The number of printed pages for 2005 was 2791. The assignment of previously accepted papers by the previous editor for 2006 (based on 2005 acceptances) followed the Fat Issue Model of 4 issues per year at roughly 550 pages per issue. Specifically, in printed pages AAP 16(1) 474 pages; AAP 16(2) 584 pages; AAP 16(3) projected 662 pages; AAP 16(4) projected 656 pages.

However, a Lean Issue Model, defined by 6 issues of say 450 pages each per year, was also proposed by the previous editor in an effort to accommodate local fluctuations around a presumed steady state rate of 2700 pages per year, and to reduce by a month the maximal waiting time from acceptance to appearance (historically 5 mos.). Based on the current backlog the journal is operating on a projected 6 to 7 months from acceptance to appearance. Apparently the Lean Issue Model continues to be a reasonable proposal to consider. There are also other issues such as impact on mailing costs to be evaluated.

The EJMS has been effective, thanks in a very large part to the dedicated and prompt support provided by Mattson Publishing Services. The EJMS



regularly requires prompt expert intervention on behalf of editors, associate editors and authors to be effective. Continued flexibility to achieve a personalized EJMS design by VTEX for IMS is also an essential aspect of the system.

While it is typically a relatively simple matter to distinguish submissions to IMS journals in areas of statistics versus probability, the line between probability and applied probability is often more vague. To a certain extent this is natural state and should be accommodated by a first priority of both AAP and AOP to be premier journals in probability. However some effort to evolve an editorial policy statement which may better help authors in the selection of an appropriate IMS journal is underway and will be further discussed with members of the editorial board at the 2006 IMS summer meeting in Rio. This could also eventually impact the number of issues model discussed above.

The current Editorial Board consists of second term continuations by: Paul Dupuis, Maury Bramson, Marco Frittelli, Claudia Klüppelberg, Russell Lyons, Leonid Mytnik, Timo Seppäläinen, Denis Talay, and Ofer Zeitouni. New (first term) members are: Jinho Baik, Rabi Bhattacharya, Brigitte Chauvin, Rick Durrett, Luc Devroye, James Allen Fill, Jean-Pierre Fouque, Charles Newman, George Papanicolaou, Vladas Sidoravicius, Stanley Sawyer, Qi-Man Shao, Jason Ross Schweinsberg, Prasad Tetali, Enrique Thomann, and Ruth Williams. The success of the journal is mostly in the hands of this very hard working and dedicated group of individuals and their anonymous referees who continue to do an outstanding job for the IMS and the larger profession.

Ed Waymire

Annals of Probability  
REPORT OF THE CURRENT EDITOR

In the period January 1, 2006 to May 31, 2006 the Annals of Probability received 122 submissions. The current status of these is the following: 5 accepted, 7 revision requested (with the expectation of eventual acceptance, 42 rejected, 2 withdrawn, and 56 still in the review process. Because of the increased number of pages and backlog of the Annals, an effort has been made to be more selective. When possible, we have tried to return papers quickly that we know will not be accepted. This includes some papers of sufficient quality that are more appropriate for Annals of Applied Probability or Annals of Statistics.

The submission and correspondence functions of the journal are handled now by the Electronic Journal Management Systems (EJMS). There has been a learning process caused by a new system and a new editor starting at the same time, but I believe that the system is working well now (although there certainly is room for change, suggestions, and improvements). Unlike previous editors, I have not been using any editorial assistance other than that provided by the system.

Greg Lawler, Editor  
2006-2008

REPORT OF THE OUTGOING EDITOR

The Annals of Probability received 221 submissions in 2005. All but 15 have been reviewed. Following is a summary of editorial decisions reached:

Accept	64
Reject	134
Revise	6
Withdrawn	2
In Review	15

The current backlog is about 6 issues (1 year). Articles accepted by the new editor, Greg Lawler, will begin to appear in the third issue of 2007.

Steve Lalley, Editor  
2003-2005

## **The Annals of Statistics, 2006 Annual Report**

Morris L. Eaton and Jianqing Fan

**SUBMISSIONS:** Submissions were up in 2005 and high relative to the historical norm: 343 manuscripts were received (in comparisons with 292, 274, 319, 362, 323 in 2000, 2001, 2002, 2003, and 2004, respectively). For the first four months of 2006, there are 128 submissions (higher than 105 submissions during the same period in 2005). Our editorial policy continues to emphasize that The Annals of Statistics aims at publishing research papers of highest quality reflecting the many facets of contemporary statistics, including all mathematical, methodological, computational and interdisciplinary work.

**ACCEPTANCE RATE:** During 2005, we have made 328 acceptance or rejection decisions on all live manuscripts. The acceptance rate among these 328 final decisions in 2005 is about 29% percent. This is compared with 30% for the manuscripts submitted during 2004. The historical acceptance rate for the Annals has been about 28-30 percent.

**BACKLOG:** During 2005, we filled volumes six to seven months in advance of the official volume date, and we are thereby achieving on-time production. The sizes of recent volumes are 1998(2469), 1999(2080), 2000(1795), 2001 (1811), 2002(1811), 2003(2095), 2004(2775), 2005(2999). There is a current backlog of about 2.5 issues. We expect to print between 3000-3100 pages in 2006. The Editors, Morris L. Eaton and Jianqing Fan, have made conscientious efforts to ask authors to streamline manuscripts before they are accepted.

**PAGE REQUEST:** The current 3000 page allotment will probably be sufficient as the allotment for 2007. The size of the backlog has been relatively stationary for the last year.

**WEBPAGE:** The Annals web page is now being maintained by Laila Hakem in connection with all the rest of the IMS website. It contains features, such as tables of contents of back issues, up-to-date lists of accepted papers, linking of accepted papers to authors' URLs, and editorial policies. We also provide links to pages of Associate Editors, as available. Another web site for administrative use of the editorial board is now being maintained by Charles Geyer along with the Editorial Coordinators Sarah Johnson Sexton and Mary Beth Falke.

**REVIEW TIMES:** About 83% percent of submissions receive an initial decision within 10 months, compared to 91% in 2004. The quartiles of the distribution of the initial decision times for manuscripts submitted in 2005 are 40 days, 180 and 255 days. They are slightly longer than the figures for the year 2004, which are 70 days, 165 days and 230 days. However

we do note that the upper tail has become somewhat heavier For those interested in finer details of the review times, the attached graph gives an estimate ‘survival curve’ of initial decision times. We continue to make efforts to shorten the upper tail of this distribution, with the cooperation of our board of Associate Editors and Referees.

ACKNOWLEDGEMENTS: The Editors, Morris L. Eaton and Jianqing Fan, would like to thank our Editorial Coordinators and Editorial Board for all their help that makes the journal running smoothly. In addition, our Editorial Coordinators, Sarah Johnson Sexton and Mary Beth Falke, and our Editorial Board, Yacine Aït-Sahalia, Peter Bühlmann, T. Tony Cai, Ming-Yen Cheng, Dorota M. Dabrowska, Rainer Dahlhaus, Anirban DasGupta, Holger Dette, Holger Drees, Lutz Dümbgen, Sam Efromovich, Charles J. Geyer, Subhashis Ghosal, Irene Gijbels, Peter Hall, Xuming He, James P. Holbert, Vladimir Koltchinskii, Michael R. Kosorok, Jan T. A. Koster, Bing Li, Jun Liu, Enno Mammen, Adam Martinsek, Thomas Mathew, Rahul Mukerjee , Hans-Georg Müller, Per A. Mykland, Michael D. Perlman, Dominique Picard, Donald St. P. Richards, Christian P. Robert, Peter M. Robinson, Qi-Man Shao, Xiaotong Shen, Vladimir Spokoiny, Michael Stein, Sara van de Geer , Mark van der Laan, Larry Wasserman , Michael Wolf, Qiwei Yao, Yi-Ching Yao, Bin Yu, Ruben H. Zamar, are currently providing the support and dedication that is essential to the success of The Annals of Statistics.

## ***IMS Bulletin: Report to Council 2005***

Bernard Silverman and Tati Howell have edited 34 issues of the *IMS Bulletin* so far. In the last four and a half years, we have implemented some significant changes in terms of production schedules – members receive their print copies earlier, and the PDF versions at [www.imstat.org/bulletin](http://www.imstat.org/bulletin) are popular, with up to 2000 downloads a month (data from May 06 website usage report [www.imstat.org/reports/](http://www.imstat.org/reports/)).

Since we do not have an ear in every Department, we rely very much on IMS members to come forward and tell us news. We invite all members, and particularly Council members, to be proactive in this area – and in offering to write longer articles. Please don't be shy! We receive occasional feedback (mostly positive) which is very welcome, but we would like to engage more with readers. If anyone has comments, suggestions or feedback, please email [bulletin@imstat.org](mailto:bulletin@imstat.org).

Across a year, there is substantial variation in the proportions of news items, meeting reports and job advertisements, and we continue to work to find a balance of interesting items for IMS members and other *Bulletin* readers.

We thank our three Associate Editors, Nicole Lazar, Susan Murphy and Ingram Olkin, who have been able to write occasional articles and pass on information about award winners. This has been very helpful. We are open to other offers of help, too, particularly from different parts of the world.

IMS Council recently approved the appointment of Xuming He (University of Illinois at Urbana-Champaign) to take on the editorship of the *Bulletin* from January 2007. Tati will remain as Assistant Editor.

*Bernard Silverman & Tati Howell*  
*June 2006*

Report to Council - IMS Lecture Note and Monograph Series

The Series continues to serve the IMS membership by providing a venue for lecture notes, monographs, conference proceedings, and festschriften.

The following volumes are in production and will appear shortly:

LNMS #48 Dynamics and Stochastics: Festschrift in Honor of M.S. Keane  
(D. Denteneer, F. den Hollander, E. Verbitskiy, eds.)

LNMS #49 The Second Erich L. Lehmann Symposium: Optimality (J. Rojo, ed.)

Ten more volumes are in various stages of preparation.

It is a pleasure to thank Executive Director Elyse Gustafson for her continuing guidance and assistance in the past year.

Respectfully submitted,  
Rick Vitale

June 15, 2006

## Report of the Managing Editor—Statistics

**Paul Shaman**

During 2005 the page count for *The Annals of Statistics* totaled 2,998, and *Statistical Science* published 430 pages.

As of this writing (June 12th) final corrections are being made to the April 2006 issue of *The Annals of Statistics*, and pages are being prepared for the June 2006 issue. Lineups have been set for the remainder of the 2006 volume, and the unassigned pending articles will fill about one and-a-half more issues. On average, the length of time from acceptance of an article until it appears in print in *The Annals of Statistics* is about ten months.

As of this writing the February 2006 issue of *Statistical Science* is ready for mailing, and pages are in production for the May 2006 issue.

The August 2005 issue of *Statistical Science* contained a special section to celebrate the 50th anniversary of the publication of “How to Lie with Statistics.”

The tables following give page counts for the past two years.

### *The Annals of Statistics*

Issue	2004 printed pages	2005 printed pages
February	406	462
April	434	514
June	471	478
August	467	527
October	561	524
December	433	493
Total	2772	2998

### *Statistical Science*

Issue	2004 printed pages	2005 printed pages
February	218	110
May	168	93
August	174	111
November	177	116
Total	737	430

The February 2006, April 2006 and June 2006 issues of *The Annals of Statistics* have 558, 489 and 531 pages, respectively. The 2006 volume will have about 3050 pages.

The February 2006 issue of *Statistical Science* has 111 pages, and the 2006 volume will have about 450 pages.

For the most recent five years the size of *The Annals of Statistics* has grown from approximately 1800 pages to roughly 3000 pages. The following table gives details.

*The Annals of Statistics*

Year	Printed pages	No. of articles <sup>1</sup>
2002	1809	64
2003	2092	76
2004	2772	94
2005	2998	98
2006	3050 <sup>2</sup>	101

<sup>1</sup>A discussion paper is counted as one article, and corrections and addenda are not included in the count.

<sup>2</sup>Estimated

For the years 2002-2005 the printed pages for *Statistical Science* totaled 467, 531, 737 and 430, respectively.



Annual Report - *Statistical Science*  
Edward I George, Executive Editor  
June, 2006

As of May 31, 2005, 22 manuscripts were at various stages of review or revision for *Statistical Science*. From June 1, 2005 – May 31, 2006, an additional 91 manuscripts were submitted. Of the total 114 manuscripts considered during this period, 49 were rejected, 36 are still in various stages of review or revision, and 29 have been accepted.

Month	Received	Rejected	Under review or revision	Accepted
Pre Jun05	22	5	5	12
June 05	4	1	3	
Jul 05	4	3	1	
Aug 05	10	5	5	
Sept 05	4	2	2	
Oct 05	9	6	3	
Nov 05	6	5	1	
Dec 05	6	3	3	
Jan 06	13	3	2	8
Feb 06	11	2	3	6
Mar 06	6	1	2	3
Apr 06	11	7	4	
May 06	8	6	2	
<b>Total</b>	<b>114</b>	<b>49</b>	<b>36</b>	<b>29</b>

I am pleased to report that both the quality and the quantity of unsolicited submissions to *Statistical Science* appear to be steadily increasing. We currently have a backlog of more than three issues. A special issue on Statistics in Electronic Commerce is currently in press, and a special issue on Reliability is almost complete. I anticipate that the total manuscript pages for 2006 will be around 500. Roughly the same number of manuscript pages is planned for 2007.

Although it took some getting used to, the EJMS system has begun to fulfill its promise. For this, I am very grateful to both Mattson Publishing Services and VTEX, who have been very responsive to my needs. With only two exceptions, all submissions to *Statistical Science* since June 1, 2005 have come through this system. The vast majority of interactions between me and the Editors and between me and the authors now occur within the system. However, I don't believe there is much, if any, interaction within the system between the Editors and referees, not that this is a problem. I am hopeful that as we go forward, EJMS will be enhanced to provide more record keeping summaries.

Finally I would like to gratefully acknowledge the efforts of my Editorial Board: Roger Berger, Alicia Carriquiry, Dean Foster, Sallie Keller-McNulty, Sally Morton, Rob McCulloch, Peter Mueller and Naisyin Wang, and the invaluable editorial support I have received from Elyse Gustafson, Patrick Kelly, Geri Mattson and Paul Shaman.

To: IMS Council  
From: Chris Burdzy, IMS Web editor  
Re: annual report

I have assumed the function of the IMS Web editor a few months ago. The previous Web editor, Hermant Ishwaran and the previous Webmaster, Arti Ishwaran, did a tremendous job of developing a very substantial and informative Web site for the IMS. The current enormous size and complexity of the Web site are the results of the rapid growth over the last few years. The size and complexity of the current Web site also necessitate its fundamental overhaul.

For the last few months, the current Web site did not undergo any substantial changes. All my energy has been devoted to the creation of the new Web site, with contents mostly based on the current Web site.

The technical development is done by Andrzej Tyburczyk, based in Lublin, Poland. He is a computer programmer whom I have known for many years. One of the reasons for contracting in Poland is that the price for the main development of the Web site is \$5,000, about one quarter of what it would cost in the US.

Other people involved in the development process are Laila Hakim, our current Webmaster (Arti Ishwaran resigned citing personal reasons), Elyse Gustafson providing crucial advice on many issues, Jim Pitman who envisioned the idea of developing database(s) for the IMS Web site and other electronic services, and Tati Howell working on the artistic aspects of the Web site.

On the technical side, the new Web site will be based on open source software - PHP, Mysql and [Mambo](#).

Some of the major parts of the project finished so far, in addition to the basic structure, include the creation of the databases for meetings and jobs.

Some of the modifications that still have to be done include:

- (i) rearrangement of the content, to make it more clear (to be done by Chris Burdzy)
- (ii) fine tuning of the artistic appearance (to be done by Tati Howell)
- (iii) development of a bookstore, to replace the current scattered pages that sell publications and memberships (to be done by Andrzej Tyburczyk).

The bookstore development will cost some extra money and will be the most challenging part of the development on the technical side.

Realistically, the new Web site will not go online for a few months. Comments and suggestions are welcome from all interested people, including members of the Council.

Respectfully,  
Chris Burdzy

## PROGRAM SECRETARY'S REPORT: AUGUST 1, 2005 – AUGUST 4, 2006

### GENERAL COMMENTS

After considerable discussion, the Executive committee decided to terminate the Minimeeting Program. The number of applications meeting the stated criteria (at most two days, informal workshop type atmosphere, primarily local participants) was very small. Rejecting most of the applications did not appear to serve the IMS or its members.

The IMS supports meetings through sponsorship and co-sponsorship. Applications for (co)sponsorship and a description of their respective benefits are available on the IMS web page. By decision of Council last year, applications for co-sponsorship are evaluated by a three person committee, currently consisting of Tom DiCiccio, Jun Liu and Erwin Bolthausen, plus the PS.

Elyse and the PS are soliciting co-sponsorship applications from appropriate meetings advertised in the IMS Bulletin and the e-Bulletin. This has led to a significant increase in the number of co-sponsored meetings.

The PS is working with officials from ENAR on a draft policy concerning the advertisement and scheduling of Medallion Lectures, and the responsibilities of the IMS PC at ENAR. We hope to establish a similar agreement with WNAR.

The Wald, Rietz and Medallion Lecturers appointed this year have all agreed to speak at their respective meetings.

### SPONSORED MEETINGS

2006 ENAR Spring Meeting

March 26-29, 2006

Hyatt Regency, Tampa, Florida, USA

IMS Co Program Chairs: Michael Kosorok and Jason Fine (UW Madison)

2006 WNAR/IMS Western Regional Meeting

June 27-30, 2006

Flagstaff, AZ, USA

IMS Program Chair: Wolfgang Polonik

31st Conference on Stochastic Processes and their Applications

July 17-21, 2006

Paris, France

IMS reps: E. Perkins, J. Pitman , P. Protter, A-s Sznitman, S. Tavaré and E. Waymire

Ninth Meeting of New Researchers in Statistics and Probability

August 1-5, 2006

University of Washington, Seattle, WA

Co-chairs: Peter Craigmile, Peter Hoff

IMS Annual Meeting and X Brazilian School of Probability (XEBP)

July 30 - August 4, 2006

Instituto Nacional de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, Brazil

Program Chairs: Robert Adler and Steve Lalley (Probability); Sara van de Geer and Guenther Walther (Statistics)

Contributed Papers Chair: Marten Wegkamp

Local Chair: Maria Eulália Vares

#### CO-SPONSORED MEETINGS

Nonparametric Models for Complex Biological Data

August 15-17, 2005

Davis, California, USA

IMS Reps: Jianqing Fan, Hans-George Mueller and Chunming Zhan

A Conference on Nonparametric Inference and Probability with Applications to Science

September 24-25, 2005

Ann Arbor, Michigan, USA

IMS Reps: Jiayang Sun and Bob Keener

Workshop on Stochastic Methods in Game Theory

September 24-October 2, 2005

Centro Majorana, Erice, Italy

IMS Rep: Marco Scarsini

25th Annual Conference of the Indian Society for Probability and Statistics

December 28-30, 2005

Bangalore University and Indian Statistical Institute, Bangalore Centre, India

IMS Reps: H.J. Vaman and S.M. Manjunath.

Frankfurter Stochastik-Tage / German Open Conference on Probability and Statistics

March 14-17, 2006

Goethe-University Frankfurt/Main

IMS Reps: Norbert Henze, Arnold Janssen, Christine Mueller, Axel Munk, Rainer Schwabe and Anton Wakolbinger

Special day on Markov Processes in honor of Erhan Cinlar's 65th birthday  
Wednesday March 22, 2006  
Princeton University, Princeton, New Jersey  
IMS Representative: Chris Burdzy

2006 Seminar on Stochastic Processes, SSP  
March 23-25, 2006  
Princeton University, Princeton, New Jersey  
IMS Representative: Chris Burdzy

Classification Society of North America 2006 Meeting on Network Data Analysis and Data Mining: Applications in Biology, Computer Science, Intrusion Detection, and Other areas  
May 10-13, 2006  
DIMACS, CoRE Building, Rutgers University, Piscataway, New Jersey  
IMS Representative: David Banks.

Workshop on Probabilistic Symmetries and Their Applications  
May 15-17, 2006  
University of Ottawa, Canada  
IMS Representatives: Gail Ivanoff and Raluca Balan

Third International Workshop in Applied Probability (IWAP 2006)  
May 15-18, 2006  
University of Connecticut, Storrs, Connecticut  
IMS Representative: Joseph Glaz

Workshop on Frontiers of Statistics  
May 18-20, 2006  
Princeton, NJ  
IMS Reps: Jianqing Fan, Luisa Fernholz and Hira Koul

International Conference on Statistical Models for Biomedical and Technical Systems  
May 29-31, 2006  
Limassol, Cyprus  
IMS Rep: Filia Vonta

Joint Summer Research Conferences (AMS-IMS-SIAM)  
June 4-29, 2006  
Snowbird Resort, Utah, USA

Statistical Challenges in Modern Astronomy IV  
June 12-15, 2006

Penn State University, Pennsylvania  
IMS Rep: G. Jogesh Babu

The Fifteenth International Workshop on Matrices and Statistics  
June 13-17, 2006  
Uppsala, Sweden  
IMS Rep: Dietrich von Rosen

Second Cornell Probability Summer School  
June 26-July 7, 2006  
Ithaca, New York, USA  
IMS Rep: Rick Durrett

International Conference on Frontiers in Statistics, Biostatistics and Bioinformatics  
July 7 - 8, 2006  
Northeast Normal University, Changchun, China  
IMS Representative: Samuel Kou

Markov Processes and Related Topics: A conference in honor of Tom Kurtz on his 65th  
birthday  
July 10 - 13, 2006  
University of Wisconsin, Madison, Wisconsin  
IMS Representative: Stew Ethier

**Report of the Treasurer  
Supplemental 2006  
Jiayang Sun**

**MEMBERSHIP AND SUBSCRIPTION DATA**

Table 1 shows the census data for 2006 memberships through 6/30/2006. Chart 1 and 2 show data for members in line and pie format. Table 1 shows subscription data for members and non-members. Chart 3 shows trends for non-member (institutional) subscriptions.

Membership continues to show upward movement. Although, the year is only half over, we can infer that membership will increase again in 2006. Subscription to print journals is decreasing as more members opt to use the gratis electronic access. This development saves the IMS funds since journals are priced at lower than cost rates for members.

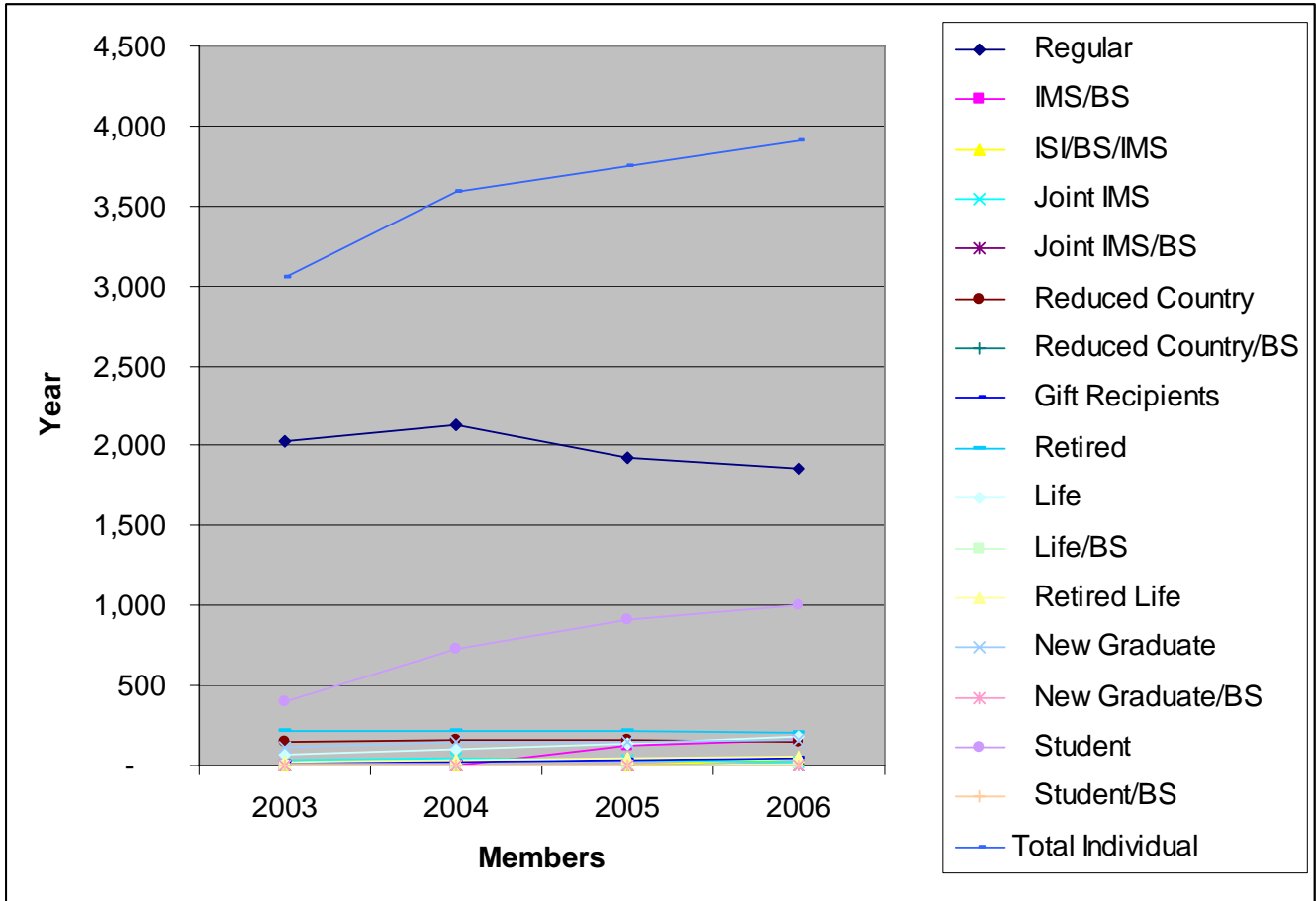
Non-member (institutional) subscriptions are showing a decrease as libraries are consolidating print collections. Libraries at large universities are moving towards keeping one print copy on campus while utilizing electronic access for the remaining campus libraries.

Not surprisingly, electronic subscriptions for members and institutions continue to rise.

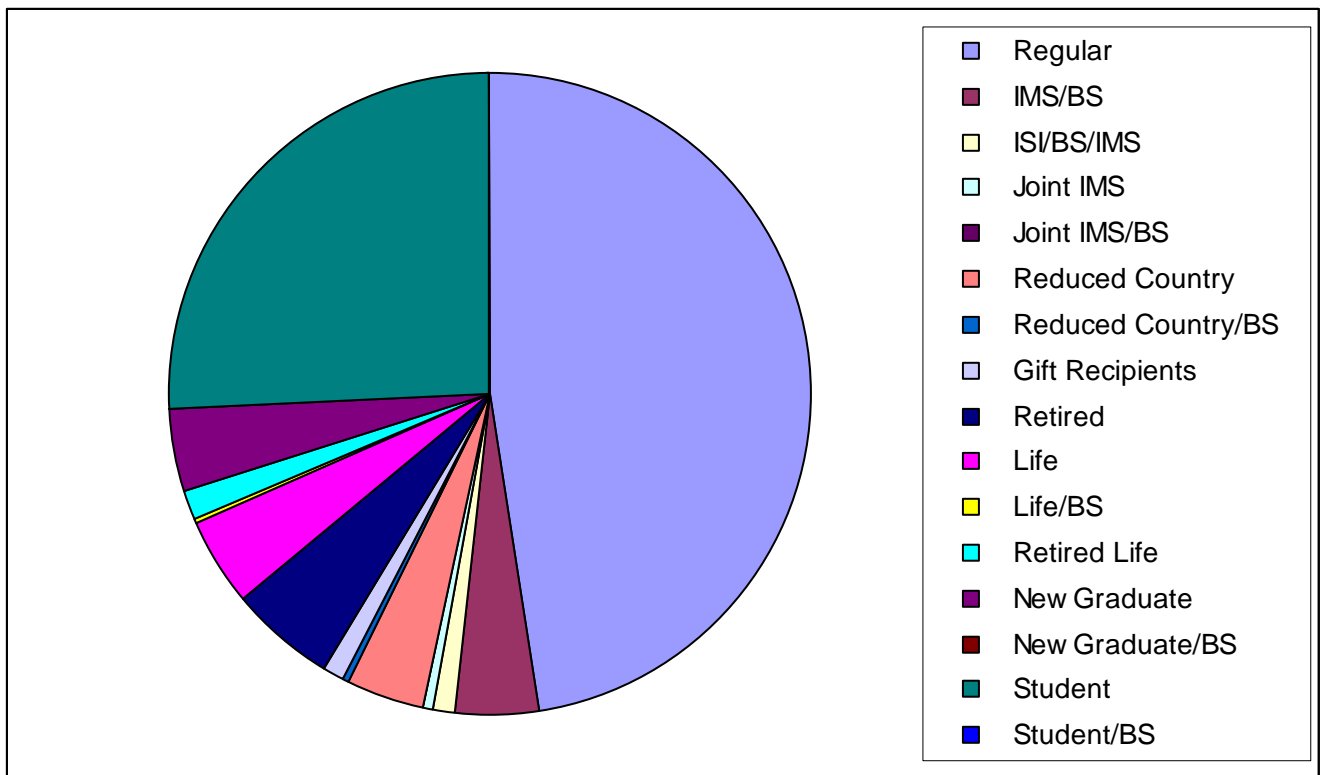
**Table 1. 2006 Member Data (6/30/2006)**

<b>MEMBERS</b>	<b>6/30/2003</b>	<b>6/30/2004</b>	<b>6/30/2005</b>	<b>6/30/2006</b>
Regular	2,028	2,130	1,928	1,852
IMS/BS	na	na	130	165
ISI/BS/IMS	na	na	na	43
Joint IMS	36	41	39	22
Joint IMS/BS	na	na	na	3
Reduced Country	152	162	154	147
Reduced Country/BS	na	na	na	9
Gift Recipients	22	25	38	45
Retired	218	222	214	206
Life	69	97	134	177
Life/BS	na	na	5	6
Retired Life	18	31	46	58
New Graduate	118	145	145	160
New Graduate/BS	na	na	na	5
Student	395	731	910	1,000
Student/BS	na	na	na	5
Total Individual	3,056	3,584	3,743	3,903
Corporate	-	1	-	2
Institutional	92	100	91	100
Total	3,148	3,685	3,834	4,005

**Chart 1. Individual Membership Trends**



**Chart 2. Membership distribution.**

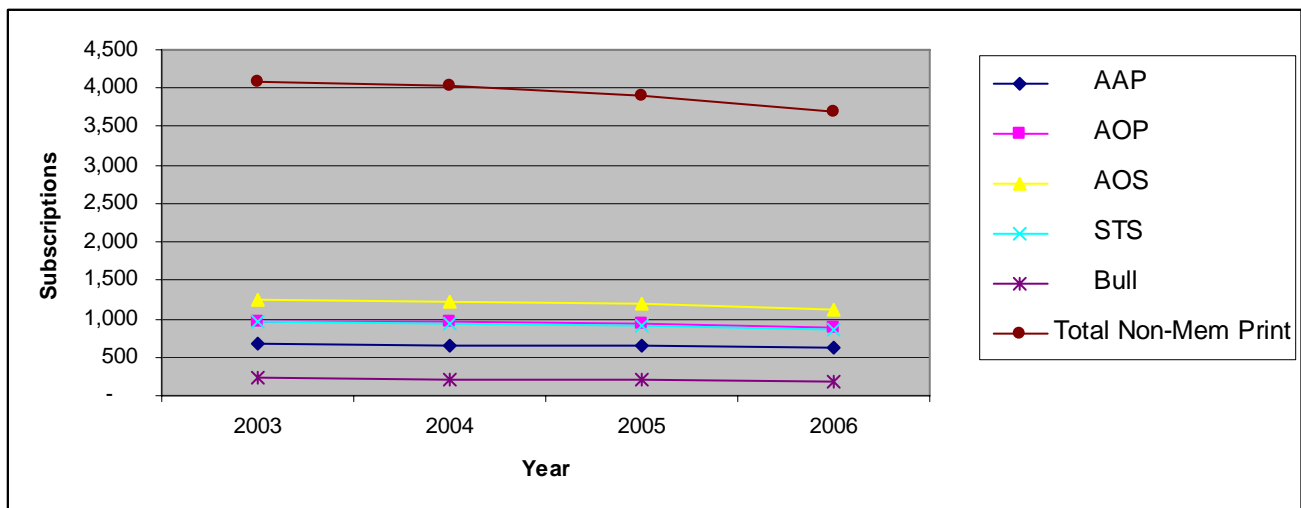




**Table 2. Member and Non-Member Subscription Data (6/30/2006)**

<b>PRINT SUBSCRIPTIONS</b>		<b>6/30/2003</b>	<b>6/30/2004</b>	<b>6/30/2005</b>	<b>6/30/2006</b>
Members					
AAP		757	807	768	745
AOP		817	858	798	763
AOS		1,698	1,828	1,769	1,667
STS		2,412	2,567	2,455	2,325
Total Member Print		5,684	6,060	5,790	5,500
Non-Member & Institutions					
AAP		665	658	645	634
AOP		961	975	946	873
AOS		1,240	1,233	1,204	1,123
STS		974	943	912	862
Bull		231	210	199	192
Total Non-Mem Print		4,071	4,019	3,906	3,684
Total Print Subs					
AAP		1,383	1,465	1,413	1,379
AOP		1,725	1,833	1,744	1,636
AOS		2,864	3,061	2,973	2,790
STS		3,294	3,510	3,367	3,187
Bull		3,331	3,996	4,124	4,299
Total Print Subs		12,597	13,865	13,621	13,291
<b>ELECTRONIC SUBSCRIPTIONS</b>					
Members					
AAP		415	686	827	934
AOP		453	711	852	950
AOS		541	941	1,166	1,333
STS		330	877	1,076	1,243
Total Member Electronic		1,739	3,215	3,921	4,460
Non-Member Institutions					
AAP		na	486	496	572
AOP		na	694	700	790
AOS		na	816	817	932
STS		na	633	643	728
Total Inst Electronic		na	2,629	2,656	3,022

**Chart 3. Non-Member (Institutions) Print Subscriptions Trends**



## **FINANCIALS**

The IMS is taking actions to increase income and decrease expenses. We continue to adjust institutional rates to accommodate the consolidation of subscriptions by universities. Because membership continues to increase, we are gaining income in this area. We are seeing many members take advantage of our new membership opportunities such as, new graduate, life and Bernoulli joint memberships.

On the expense side, we have instituted the central editorial office to help decrease editorial expenses. There will be one editorial office serving 6 journals (AAP, AOP, AOS, STS, PS and the new AoAS). Most actions are automated electronically. The system implementation will be complete in FY2007. We expect to save a minimum of \$50K once all journals are on the system.

Four years ago, the IMS opened full electronic access to all IMS members regardless of their print subscriptions. This action added a strong benefit for our members while at the same time decreasing our expenses. As member print subscriptions decrease we are saving on publication expenses. Members currently pay less than the variable cost on print journals.

Expect to see a full Treasurer's report in the March 2007 IMS Bulletin. The report will include final 2006 membership and subscription data as well as the FY05 audit and financial reports.

## Annual Report of the IMS Treasurer

### INTRODUCTION

This report details membership and subscription data for calendar year end 2005. In addition, it reviews the FY2005 (July 1, 2004 – June 30, 2005) financial statements. I am proud to announce, for the sixth year in a row the IMS experienced another increase in total membership. We have many members taking advantage of the new membership options that have been introduced over the last five years. For 2006, we have added a new joint IMS/ISI/Bernoulli Society membership which allows members to receive 25% off membership when they join each society.

Several years ago, the IMS Executive Committee and Council decided to invest more funds back into our membership. Several programs reflect this new philosophy including:

- open access ArXiv placement of all IMS articles;
- free student membership and print journal;
- reduced dues for new graduates;
- gratis electronic access to all journals past and present for all members;
- discounts for on-time renewal;
- child care for those attending the IMS Annual Meeting; and
- travel awards for students and new graduates.

The financial status of the Institute continues to be strong and stable. Details of the events of the past year, membership and subscription data, sales data and a detailed analysis of the financial statement for FY 2005 are given below.

### Dues and Subscriptions Office.

The IMS continues our agreement with the Federation for Societies in Experimental Biology (FASEB) to continue handling all dues and subscription processing. The IMS relationship with FASEB has been in place since 2000. We have found increased efficiencies and economies of scale that allow us to invest more funds toward membership benefits rather than administration.

### Societal Office.

Elyse Gustafson is in her ninth year as our Executive Director. She continues to handle all societal issues from her office in Cleveland, Ohio. Elyse will provide a full report on activities from her office in an upcoming IMS Bulletin. Please be sure to read it.

### Publications and Web.

*Journals.* IMS journals are the core of our mission. The annals continue to be top tier in the field. In 2004, the IMS began placing all IMS articles on the open access ArXiv (see <http://www.imstat.org/publications/arxiv.html> for more information).

*IMS Journal Editors.* Welcome to Ed Waymire, Editor, *Annals of Applied Probability* and Greg Lawler, *Annals of Probability*. We owe special thanks to Robert Adler and Steve Lalley who have completed their terms and have served the IMS with great dedication.

*Electronic Access.* All IMS members receive electronic access to all IMS journals (1996 to date) through Project Euclid. In addition, members whose organizations do not subscribe to JSTOR can receive individual access to all IMS journals (1930-2001) via JSTOR. For more information see <http://www.imstat.org/publications/eaccess.htm>. Journals on Project Euclid older than 3 years are fully open to the public.

*IMS Electronic Journals.* The IMS currently sponsors three open access electronic journals, Electronic Journal of Probability, Electronic Communications in Probability and Probability Surveys. All three journals can be accessed via the IMS web page <http://www.imstat.org>. We welcome the new editors for EJP and ECP, Andreas Greven and David Nualart (respectively).

*IMS Lecture Notes - Monograph Series and NSF-CBMS Regional Conference Series.* Since my last report there have not been any new volumes in either of these series. We expect at least one LNMS volume in 2006.

### IMS Meetings and Awards.

During FY 2005, the IMS granted a total of \$25,470 to support students and new graduates and those in developing countries. \$21,041 was granted to 2005 Laha Travel Award Recipients and \$4,429 was granted to support lecturers for the IMS Visiting Lecturer in Statistics Program.

### MEMBERSHIP DATA

Total membership in the Institute as of December 31, 2005 was up 6.7% from December 31, 2004. Table 1 presents the distribution of memberships by category for the last several years. Individual membership decreased as life membership increased accounting for 72 members opting to upgrade to life membership.

*Breakdown of Member Categories.* Among the individual members for 2005, a total of 50 are Gift members (31 last year), 41 are joint members (40 last year), 219 are retired (225 last year) and 166 are reduced rates (175 last year) and the remaining 2,206 are regular members (2,275 last year). Within the Life membership category, 50 are retired life members (31 last year) and 155 are regular life members (102 last year).

*Geographic Distribution of Members.* The IMS membership is currently distributed as such, 64% USA, 17% Europe, 6% East Asia, 5% Canada, 2% West and Central Asia, 2% Australia and New Zealand, 1% South America, 1% Middle East, 1% Central America and the Caribbean, and 1% Africa.

**TABLE 1**  
Distribution of Memberships by Category  
Calendar Year Data (Jan-Dec)

	1998	1999	2000	2001	2002	2003	2004	2005	% change
Individual	3097	2898	2777	2820	2758	2744	2746	2682	-2.33%
Life	11	11	10	8	51	115	133	205	54.14%
New Graduate	na	na	na	93	131	122	165	187	13.33%
Student	182	228	478	395	496	707	971	1224	26.06%
Organizational	98	100	96	94	98	102	107	100	-6.54%
Total	3388	3237	3361	3410	3534	3790	4122	4398	6.70%

*Selection of Journals by Members.* Print subscriptions by members were up in 2005; this can be attributed to the program instituted in 2004 providing one free print journal to all students. The increases are attributed to students, whereas individual members are opting to decrease print subscriptions while enjoying free electronic access to all journals.

Revenue from all Institute member dues and journal subscriptions amounted to \$355,253 for the fiscal year ending June 30, 2005, down slightly from \$357,657 in FY 2004. This is attributed to decreased print subscriptions.

**TABLE 2**  
Distribution of Journal Selections by Members  
Calendar Year Data (Jan-Dec)

<b>PRINT</b>									
Members	1998	1999	2000	2001	2002	2003	2004	2005	% change
AAP	1,048	911	911	902	865	844	800	870	8.75%
AOP	1,139	1,007	918	952	918	910	907	877	-3.31%
AOS	2,107	1,950	1,904	1,992	1,949	1,917	1,987	2,053	3.32%
STS	2,691	2,672	2,661	2,707	2,778	2,846	2,750	2,765	0.55%
Total Member Print	6,985	6,540	6,394	6,553	6,510	6,517	6,444	6,565	1.88%
<b>ELECTRONIC</b>									
Members	1998	1999	2000	2001	2002	2003	2004	2005	% change
AAP	na	na	na	na	363	715	820	889	8.41%
AOP	na	na	na	na	411	693	791	902	14.03%
AOS	na	na	na	na	482	943	1,112	1,262	13.49%
STS	na	na	na	na	295	877	1,023	1,168	14.17%
Total Member Electronic	na	na	na	na	1,551	3,228	3,746	4,221	12.68%

#### NON-MEMBER SUBSCRIPTION DATA

Table 3 presents comparative subscription data for non-members to each of our scientific journals for 2005 and previous years. All journals experienced decreases in print subscriptions in 2005, while electronic subscriptions continue up. Revenue from all non-member subscriptions was \$713,767 for the fiscal year ending June 30, 2005, up from \$655,159 for the FY 2004. The increase is due to increased subscription rates for 2005. Approximately 60% of the non-member subscribers to IMS journals are in USA and Canada, with the remaining subscribers distributed throughout the world.

**TABLE 3**  
Distribution of Journal Selections by Non-Member Subscribers  
Calendar Year Data (Jan-Dec)

**PRINT**

Non-Member Institutions	1998	1999	2000	2001	2002	2003	2004	2005	% change
AAP	777	799	779	680	690	716	675	659	-2.37%
AOP	1,148	1,127	1,121	983	1,001	1,034	1,001	974	-2.70%
AOS	1,512	1,481	1,454	1,305	1,320	1,342	1,268	1,233	-2.76%
STS	1,180	1,156	1,258	1,068	1,041	1,064	976	949	-2.77%
Bull	249	284	320	259	267	229	222	207	-6.76%
Total Non-Mem Inst Print	4,866	4,847	4,932	4,295	4,319	4,385	4,142	4,022	-2.90%

**ELECTRONIC**

Non-Member Institutions	1998	1999	2000	2001	2002	2003	2004	2005	% change
AAP	na	na	na	na	na	363	480	514	7.08%
AOP	na	na	na	na	na	520	684	713	4.24%
AOS	na	na	na	na	na	593	800	857	7.13%
STS	na	na	na	na	na	459	635	677	6.61%
Total Non-Mem Inst Electronic	na	na	na	na	na	1,935	2,599	2,761	6.23%

**SALES DATA**

There was one new volume in the *NSF-CBMS Regional Conference Series in Probability and Statistics* in FY 2005. In FY 2005, total revenue from this Series was \$7,527, stable compared to \$7,599 in FY 2004. Table 4 shows summary data on sales from the NSF-CBMS Regional Conference Series.

Three new volumes in the *Lecture Notes - Monograph Series* were published in FY 2005. Table 5 presents sales data for Volumes 1-47 of this Series. Total revenue from the Series decreased to \$24,303 in FY 2005 from \$30,540 in FY 2004.

**TABLE 4**  
Sales from the NSF-CBMS Regional Conference Series  
Fiscal Year Data (July 1-June 30)

Vol	Short Title	Year Published	to 2000	2001	2002	2003	2004	2005	TOTAL
1	Group Invariance	1989	854	11	11	7	19	6	908
2	Empirical Pro	1990	1,033	34	41	32	64	21	1,225
3	Stochastic Curve	1991	620	8	19	10	10	7	674
4	Higher Order	1994	412	17	15	24	26	11	505
5	Mixture Models	1995	713	108	40	53	76	30	1,020
6	Genetic Data	2000		306	194	136	75	28	739
7	Linear Mixed Models	2003				45	124	70	239
8	Longitudinal	2004						155	155
Total			3,632	484	320	307	394	328	5,465

**TABLE 5**  
LNMS Sales  
Fiscal Year Data (July 1-June 30)

Vol	Short Title	Date	to 2000	2001	2002	2003	2004	2005	Total Sales
1	Essays	1981	967	4	6	4	8	4	993
2	Survival	1983	1,260	7	7	5	9	6	1,294
3	Empirical	1983	1,059	11	9	8	13	11	1,111
4	Zonal	1984	695	11	10	6	11	7	740
5	Inequalities	1984	1,267	11	13	5	12	12	1,320
6	Likelihood	1984	2,104	37	31	50	34	30	2,286
7	Approximate	1987	1,119	17	11	12	17	21	1,197
8	Adaptive	1987	560	5	7	3	3	10	588
9	Fundamentals	1987	1,374	23	25	48	22	17	1,509
10	Differential	1988	1,434	23	29	27	19	20	1,552
11	Group	1988	1,941	45	44	33	52	36	2,151
12	Continuity	1990	746	11	16	18	20	8	819
13	Small	1990	1,031	17	17	13	10	16	1,104
14	Invariant	1991	280	9	10	7	14	10	330
15	Analytic	1991	429	7	6	4	7	5	458
16	Topics	1991	427	7	5	5	6	5	455
17	Current	1992	387	7	4	6	3	7	414
18	Selected	1992	306	5	3	2	1	3	320
19	Stochastic	1992	289	9	6	6	6	2	318
20	Spatial	1992	501	5	8	8	7	4	533
21	Weighted	1992	428	8	11	8	2	5	462
22	Stochastic	1993	305	8	6	5	9	7	340
23	Change-point	1994	311	16	10	10	10	7	364
24	Multivariate	1994	301	13	9	9	10	11	353
25	Adaptive Des	1995	225	13	10	10	4	1	263
26	Stoch Diff Equ	1996	269	14	12	11	12	8	326
27	Anal Cens Data	1996	265	12	11	6	11	8	313
28	Fixed Marginals	1997	250	20	15	13	14	6	318
29	Bayes Robust	1997	285	17	14	12	9	9	346
30	Game Theory	1996	305	12	11	8	4	3	343
31	L1 Statistical	1997	243	10	7	4	3	3	270
32	Selected Pro	1998	249	16	10	7	7	4	293
33	Mol Biology	1999	192	117	57	11	9	6	392
34	Exp Design	1998	160	51	20	8	8	7	254
35	Ferguson	2000		79	75	22	14	6	196
36	State of the Art	2000		2	105	21	10	5	143
37	Inference	2001			61	38	12	5	116
38	Model Selection	2001			121	153	61	38	373
39	Bahadur	2002				211	61	45	317
40	Speed	2003				47	60	12	119
41	Bhattacharya	2003				26	66	13	105
42	van Eeden	2003					80	14	94
43	Hall	2003					79	16	95
44	1st Lehmann	2004					58	34	92
45	Rubin	2004						27	27
46	Stein's Method	2004						39	39
47	Multiple Comp	2004						30	30
<b>TOTAL</b>			21,964	679	832	910	887	603	25,875

## FINANCIAL OVERVIEW

This is a detailed analysis of the Financial Statement for FY 2005, which is presented in this issue of the *IMS Bulletin*, following this Treasurer's Report. Comparisons are always with FY 2004. The overall picture of the financial status of the Institute is strong and stable. Per the auditor's report, in FY 2005 we experienced an increase in total assets of \$66,954. The increase is due to the increase in life members and increase in value of joint ventures. The IMS experienced a decrease in unrestricted net assets of \$115,611. The following will explain the reasons for the decrease. The IMS has strong reserves and it has been the goal of the Council to put our revenues back into services to the membership and the community. This was done on such programs as travel grants, free electronic access, pre-print posting of articles, early renewal discounts, free student members, free journals for students, reduced dues for new researchers and more. The Statement of Activities shows an increase in total revenue and in total expenses compared with FY 2004. Total revenues are lower than expenses showing a net loss.

### Revenue.

Membership dues and subscription revenues were adjusted, as in the past to prorate calendar-year revenues to fit with the Institute's fiscal year reporting. Revenues from membership dues and subscriptions are stable as compared to FY 2004. Revenues from non-member subscribers are up due to increases in subscription rates. The contributions listed in FY 2004 and FY 2005 represent donations made to the Tweedie Memorial Fund. Sales of back issues are down from FY 2004 as we now only sell three years back to decrease storage expenses. Page charges are up. Due to the voluntary nature of the contributions, the levels received tend to fluctuate. Revenue from sales of *Lecture-Notes Monograph Series* was down as only three volumes were released in the fiscal year and general sales are down. Revenue from sales of *NSF-CBMS Series* were stable. Meeting income increased as we did not handle funds for any meetings. Advertising revenues were up due to the Bulletin switching from six issues annually to ten. Offprints, royalty and other category is up as royalties from IMS's interest in JSTOR increased and IMS received income from CIS in FY05. Net earnings of joint publication ventures show a slight increase in FY 2005. The publications' (*Current Index to Statistics* and the *Journal of Computational and Graphical Statistics*) management committees have been working to address the issues facing the publications. The unrealized loss on investments is merely a line item, which shows prepaid interest and is not an *actual* loss or gain on investments. That amount should be totaled with the Investment Income line item to get a complete understanding of our gain on investments in FY 2005. Investment income is up in FY 2005 as interest rates are once again on the rise.

### Expenses.

The IMS makes a distinction between Program and General Administrative expenses in its audited reports. This is appropriate reporting for a non-profit organization and gives members a better idea of how much is being spent on actual programming (journals, meetings, etc) versus what is spent purely on administration of the Institute. I am happy to report that 95.5% (up from 95.0% last year) of your dues dollars goes directly into the program functions of the IMS.

*Discussion of Note G.* Here you will see the allocation for expenses for Program and General Administrative. Production and Editorial expenses will be discussed below in the "Discussion of Note H."

The management fee shows the expenses paid to FASEB for their dues, subscriptions and web services and is up for FY 2005 as we begin using more services from FASEB, including mailings, web hosting, email hosting and web services. Salaries are up in FY 2005 reflecting wage increases and the use of a temporary staff as needed. Mailing and shipping at the press is up from FY 2004, as postage rates increased, total pages and issues of journals increased. Scientific meeting expenses are up from FY 2004 as more travel grants were provided in FY2005. Business meeting expenses were up due to the need for travel by executive committee members to Europe and Singapore. Rent and utilities are steady. Contributions to other societies is up slightly. Postage and printing are down from FY 2004, as fewer renewals had to be sent out as people renewed on time due to discounts and agent subscription orders were batched for shipment. Computer equipment and software was up as new equipment was needed. Professional fees were up as more legal input was needed. Insurance fees are stabilizing after a few years of increases. Storage fees were up as we had more journals in storage. This number should decrease next year as fewer back issues are being kept on hand. Supplies are steady. Telephone is up. Membership drives and publicity is down as one marketing item expected for FY05 did not go out until FY06. Office expense includes bank fees and other miscellaneous expenses.

*Discussion of Note H.* Production expenses for *Annals of Statistics*, *Annals of Applied Probability* and *Statistical Science* were all up due to increases in pages or issues during FY 2005. *Annals of Probability* is stable. *LNMS* expense is stable. The *NSF-CBMS Series* had reprint expenses in FY 2005. Electronic operations for all expenses include fees for placement of our hosting of our journals on Project Euclid and metadata generation, there is an increase as one new journal, *Probability Surveys*, was added in FY 2005. Editorial expenses for *The Annals of Applied Probability* and *Annals of Probability* are stable. *Statistical Science* is down as it has been moved to a Central Editorial System, as the other journals will be as editors change. *The Annals of Statistics* is up as both a bill from previous editors came in FY 2005 although the editors had completed work more than one year prior. All editors are within their budgets for the length of their term. The *IMS Bulletin* editor is up due to moving from six issues to ten annually. Managing and production editorial expenses are up due to inflation and increased services due to increased total pages. The Web editor expenses are down as fewer services were utilized in FY 2005.

**Recommendation.**

This year we recommended an institutional subscription rate increase of ~10% for 2006. Dues and journal rates for members remain the same for 2006. Members were given a 20% discount off dues if they renewed by December 31. The 2005-2006 Council approved these recommendations at the Annual Meeting in August 2005 in Minneapolis, MB, USA.

*January 2006*

*Jiayang Sun, Treasurer*



To: ASA and IMS Boards of Directors  
From: Duncan Murdoch, Chair, CIS Management Committee  
Re: Reappointment of Abstracting Editor; Budget for CIS  
Date: May 24, 2006

## **Status report**

The Current Index to Statistics is currently housed at Carnegie Mellon University (with David James as Database Editor). Operations are running smoothly, and journal article updates are now occurring approximately quarterly. David is planning an effort over the summer to bring our book entries up to date, linking books to their reviews.

Distribution of bibliographic information has changed drastically over the lifetime of CIS. Recently it has become the norm for researchers to obtain copies of articles online, rather than by photocopy from a library. Currently CIS has no automatic links to online content. We had planned to add this over the summer of 2005, but this did not happen. Instead our efforts went towards a proposal to move CIS hosting to Cornell University Libraries. Unfortunately, this turned out to be much more expensive than we had anticipated, and the plan has been shelved. We are currently planning more modest proposals for contracting out particular aspects of the work.

In 2004, the MC agreed to allow free online access for ASA and IMS members. This is now in place from both the ASA and IMS web sites. As expected, we have seen a substantial drop in personal subscriptions due to this change, with a corresponding drop in revenue. However, costs have been lower than expected, so the net result is an increased surplus in 2005 over 2004. We expect 2006 will come in approximately on budget, i.e. with revenue and expenses approximately balanced.

In last year's budget, we planned for an open competition for web development proposals. We carried out this competition, but only had 3 submissions. We do not plan to hold a similar competition in 2007. A full report on the competition is attached.

## **Reappointment of Abstracting Editor**

The term of appointment for our Abstracting Editor, George Styan, is due to expire on December 31, 2006. Styan has expressed a willingness to extend his term by one more year, and the MC has agreed to nominate him for an additional year (i.e. to December 31, 2007). A motion to this effect is attached.

## **Search for Database Editor**

The term of appointment for our Database Editor, David James, is also due to expire on December 31, 2006. We are in the process of searching for a new editor.

## **Medium to Long Range Plans**

It is clear that CIS must add features in order to remain useful. It is no longer sufficient to give users citations of papers and expect them to find the papers themselves; we need to provide means to get to the papers, as competing bibliographic services do.

Fortunately librarians and others have developed standard schemes to facilitate this. CIS can construct simple links to brokers who will redirect them to full text versions of articles for users who are subscribers to the original journal. We are currently in the process of arranging a contract to carry out this work.

CIS will also become more relevant to its users as a trusted reference source. We are planning to implement unique article identifiers, so authors can cite a link to CIS which will give readers access to all the CIS information on an article.

These and similar initiatives require time and energy to implement. We are planning to contract out some of the development work: it is too much to expect volunteers to implement. We expect that these improvements to CIS will make it more attractive to libraries and commercial subscribers. Potentially, its market is much larger than the current subscription base: doubling or tripling our library subscriptions would be an achievable target in 3-5 years, if we offer more value to our users. On the other hand, maintaining the status quo does not appear to be sustainable in the long term.

## **2007 Budget Proposal**

In putting together the 2007 budget, we considered the following on the revenue side:

- We are concerned about competition from free services such as Google Scholar, and our own free access to ASA and IMS members, so we have not raised subscription rates.
- The forecast numbers of subscriptions were chosen to be between the 2005 actual and 2006 budgeted numbers. We expect increases in the long term as we implement and advertize new features such as linking to online content, but are not counting on this increase to be reflected in 2007 revenue.

On the expenses side:

- We have kept the Abstracting Editor's budget fixed at \$30,000. As predicted last year, the Canadian dollar has risen relative to the US dollar; our budget assumes there will not be further increases through 2007.
- We have added \$10,000 to the Database Editor's budget. The 2006 budget of \$25,000 included \$10,000 for the web development project competition which we do not plan to hold in 2007, so this actually represents a \$20,000 increase in the money available to the Database Editor. We expect to use this money to contract out specific web development projects: linking to external content, etc.

- We have deleted the budget item of \$1500 in licensing fees. It turned out that CIS is covered under ASA subscriptions for the services we needed.
- The CIS admin assistant at Harvard has been budgeted at the same amount as in 2006.
- We have increased the computing equipment budget from \$1000 to \$3000. We expect that the new Database Editor will need to purchase equipment to support the operations.
- Other items have been budgeted the same as 2006.

### **Approvals Requested**

On behalf of the MC, I would like to request that the Boards of the ASA and the IMS approve the budget and the extension of the Abstracting Editor's term.

Respectfully submitted by Duncan Murdoch, Chair, CIS Management Committee

Encl: Budget spreadsheet  
Report on web development competition

**CURRENT INDEX TO STATISTICS**

**DRAFT 2007 Budget**

	2004 <u>Actual</u>	2005 <u>Actual (adjusted)<sup>1</sup></u>	2006 <u>Budget</u>	2007 <u>Budget</u>	<u>Calculation of 2007 revenue projections</u>					
					2005 <u>Sales</u> Actual	2006 <u>Sales</u> Budgeted	2007 <u>Sales</u> Budgeted	2005 <u>Unit</u> <u>price</u>	2006 <u>Unit</u> <u>price</u>	2007 <u>Unit</u> <u>price</u>
<b>REVENUE</b>										
Sales - General	\$20	\$200	\$-	\$-						
Sales - CIS ED 4 Station License	1,000	-	-	-						
Sales - CIS ED 4 Station Upgrade	625	-	-	-						
Sales - CIS ED Personal License	4,635	-	-	-						
Sales - CIS ED Personal Upgrade	1,035	-	-	-						
Sales - CIS ED Commercial Upgrade	2,160	-	-	-						
Sales - CIS ED University License	5,400	-	-	-						
Sales - CIS ED University Upgrade	1,354	-	-	-						
CIS Web - Personal	-	-	1,950	650	-	30	10	65	65	65
CIS Web - Commercial	32,311	29,597	24,000	29,520	41	33	41	720	720	720
CIS Web - General	11,976	10,924	13,200	13,200	27	33	33	400	400	400
CIS Web - University	63,589	62,709	54,000	59,940	232	200	222	270	270	270
CIS Web - 4 Station	5,217	4,067	4,725	4,750	33	38	38	125	125	125
Royalties	93	6	-	-						
Interest - Other	853	1,428	-	-						
<b>Total Revenue</b>	<b>\$130,268</b>	<b>\$108,929</b>	<b>\$97,875</b>	<b>\$108,060</b>						
<b>EXPENSES</b>										
CIS Abstracting Editor (Styan)	\$26,000	\$28,000	\$30,000	\$30,000						
CIS Database Editor (New editor)	1,696	6,071	25,000	35,000						
Database licensing fees	-	-	1,500	-						
Computing Equipment (New editor)	2,321	-	1,000	3,000						
Subtotal - program	<u>\$30,017</u>	<u>\$34,071</u>	<u>\$57,500</u>	<u>\$68,000</u>						
CIS Admin Assistant (Harvard Med Sch)	\$13,728	\$10,975	\$14,500	\$14,500						
Software Support (Carnegie Mellon)	-	541	-	-						
Postage / Shipping Charges	67	82	200	100						
Telephone	25	0	200	500						
Storage/warehouse	-	-	500	-						
Audit/Accounting Services	4,700	730	5,000	5,000						
Bank Charges	955	735	1,000	1,000						
Bank Credit Card Fees	2,611	2,165	3,000	3,000						
Publications Management Committee	482	-	2,500	2,500						
Awards/Plaques (for retiring editor)	68	-	-	100						
Marketing/Advertising	-	-	400	500						
ASA Management Costs	<u>20,700</u>	<u>9,000</u>	<u>12,500</u>	<u>12,500</u>						
Subtotal - administrative	<u>\$43,336</u>	<u>\$24,228</u>	<u>\$39,800</u>	<u>\$39,700</u>						
<b>Total Expenses</b>	<b>\$73,353</b>	<b>\$58,298</b>	<b>\$97,300</b>	<b>\$107,700</b>						
<b>NET REVENUE (EXPENSE)</b>	<b>\$56,914</b>	<b>\$50,631</b>	<b>\$575</b>	<b>\$360</b>						
ASA share	50%	\$28,457	\$25,315	\$288	\$180					
IMS share	50%	28,457	25,315	288	180					
		<u>\$56,914</u>	<u>\$50,631</u>	<u>\$575</u>	<u>\$360</u>					

<sup>1</sup>Adjusted to include \$10,927 FY2005 reimbursement request for CIS admin assistant received after books closed. This expense will be reported in the 2006 books, but is reflected in the 2005 column of this spreadsheet for clarity. Also adjusted by DJM to re-label funding for database assistant Paul Juska as a Database Editor expense rather than an Admin Assistant expense, and to delete items with no expense historically or expected.

# Results of CIS call for web development proposals

The CIS Management Committee

April 21, 2006

In its budget submission for fiscal 2006, CIS included \$10,000 to fund a competition for web development projects. The RFP for these went out in January 2006, with a deadline for submissions of April 1, 2006. The goal of the program was to add value to CIS. Projects were to be judged by the CIS Management Committee on the criteria:

- Value to CIS of the likely outcome from the project.
- Clarity of objectives, and likelihood of success.
- Matching funds from other sources.

In all, three proposals were received, including one from one of the members of the CIS Management Committee (Jim Pitman). The other two were from Valentinas Kriauciukis and Sigitis Tolusis of VTEX, and Andrew McCallum of the University of Massachusetts. The proposals are attached to this report.

Prof. Pitman excused himself from the discussion of the merits of the proposals. The other four members of the Management Committee (Duncan Murdoch, as chair; Ed Gbur, Jim Gentle for the ASA; John Wierman for IMS) read the proposals. Murdoch discussed them by telephone with Wierman and Gbur, and received comments by email from Gentle.

All three proposals appeared to meet the criteria set by the Management Committee. We ranked the Pitman and McCallum proposals ahead of the VTEX proposal based mainly on two considerations: they were both valuable contributions unlikely to be pursued by CIS without this funding, and their

budgets requested \$5,000 each, allowing both to be funded within the \$10,000 budget for the program.

The VTEX proposal will provide a core service that CIS has identified as being a high priority, namely providing links from CIS results to other sources of information on the Internet about each article: full text, Math Reviews entries, etc. We recommend that it be pursued using other CIS funds. The full CIS Management Committee should decide how best to proceed: whether to use 2006 funds or budget it for 2007.

# CIS Proposal: BibServer development

Jim Pitman

March 31, 2006

## 1 Background

BibServer is a Python program designed to create a network of displays of bibliographic data maintained in BibTeX or logically equivalent format (XML, YAML, ... ) by contributing authors and editors. Current code is capable of providing multifaceted displays over all many kinds of bibliographic data, including navigation over sets of people as well as varied sets of bibliographic items. Following are some sample pages:

- Author Index  
<http://bibserver.berkeley.edu/cgi-bin/bibserver>
- Individual author listing with search facility  
[http://bibserver.berkeley.edu/cgi-bin/main?au=UCB\\_MATH:74](http://bibserver.berkeley.edu/cgi-bin/main?au=UCB_MATH:74)
- UC Berkeley Mathematics Faculty Listing  
[http://bibserver.berkeley.edu/cgi-bin/authors?&Source=UCB\\_MATH](http://bibserver.berkeley.edu/cgi-bin/authors?&Source=UCB_MATH)

Proposer is collaborating with a number of others, including staff at VTEX, and Nitin Borwankar, web applications developer, to modularize the bibserver code in an appropriate web application architecture for use by numerous individuals and departments.

Once the code base has stabilized and been made portable, VTEX has agreed to maintain it and use it to support the IMS Bio-bibliographies Project, which was recently approved by IMS Council. As a benefit of membership, IMS will provide technical assistance to encourage its members to create and maintain machine-readable bio-bibliographies to be used for a number of purposes:

1. for members to provide high quality easily maintained web-renderings of their biobib data, such as the bibliographies provided by BibServer
2. to assist automated acquisition of bibliographic content by Current Index of Statistics, in particular for CIS to provide author name authority in its bibliographic records (Math Reviews already does this, and has agreed in principle allow extension of their identifiers to CIS)

3. to assist editing and presentation of biobib data in memorial volumes and festschrifts
4. to assist in creation of an attractive online compilation of collected works of IMS members, such as the Dutch Cream of Science Project

In the first instance, this will be done for IMS Fellows, then expanded to broader membership. Specifically, VTEX will prepare suitable latex/bibtex/xml templates for management of biobib data, so these templates can be used routinely in preparation of memorial volumes and festschrifts. As a benefit of membership, IMS will provide a service to its members whereby VTEX would convert their biobib data from legacy formats into the latex/bibtex/xml standard.

## 2 Proposal

It is proposed to continue work on modularization of the bibserver code before transfer of the code base to VTEX to manage. Funds are requested to support this programming effort.

- Budget: \$5,000 to support coding work by Nitin Borwankar.
- Timeline: BibServer code base should be transferred to VTEX for use in the IMS biobibs project by June 2006.
- Deliverables: Properly modularized open source python code with typical MVC web application architecture capable of providing current bibserver functionality over a suitably configured MySQL database.
- Value to CIS of the outcome: Individuals and departments will have adequate software, either to manage themselves, or to access by a VTEX operated webservice, to serve their own bibliographic data to the web. Format of this data should be adequately structured to allow CIS to access this data in an automated way, with the option of adding data in the index in a semi-automated way, subject to approval by editors.
- Clarity of objectives, likelihood of success: This proposal a small component of the broader program of bibserver development, for which there is substantial support in the community, as evidenced by IMS and the Departments of Mathematics and Statistics at U.C. Berkeley.
- Matching funds from other sources: \$5,000 towards coding work has already been provided by IMS, which is planning to support VTEX work on the IMS biobibs project. It is expected that further support of customization of bibserver for use by U.C. Berkeley Mathematics department faculty and staff will be provided by the Berkeley Center for Pure and Applied Mathematics, likely around \$15,000. This will provide the administrative interface to allow faculty and staff to update their bibserver entries.



### **3 Followup**

It is hoped that CIS will make a long-term commitment to data exchange with bibserver installations managed by various departments and other organizations with bibliographic data to share, by provision of some financial support to maintain the bibserver infrastructure. A start should be provided by cooperation between CIS and IMS/VTEX on data standards for the IMS biobibs project. CIS has already agreed to allow its data to be used for IMS biobibs displays. The present proposal is a step towards facilitating data transfer in the other direction, to allow CIS to import data from sources maintained by numerous individuals and groups using bibserver software to manage their data.

# CIS enhancement with linkserver

## Proposal for CIS Management Committee

Valentinas Kriaučiukas and Sigitas Tolušis  
(valius@vtex.lt, sigitas@vtex.lt)

VTEX

March 29, 2006

### 1 Introduction

This is a proposal for the Current Index to Statistics (CIS) following Call for Proposals of the CIS Management Committee.

We propose to enrich CIS database entries with multiple links through *linkserver*, a new kind of web service, something like presented in the demo. The *linkserver* transforms one permanent standard link from a CIS item to the updatable corresponding links to MathSciNet, arXiv, Zentralblatt Math, CrossRef, etc. It allows quick and direct access to the repositories containing full text, abstracts, reviews of articles. New targets can be added in future without change of the parent link from CIS.

### 2 Requirements

1. The linkserver accepts SICI identifiers (other forms of identifiers can be added in the future).
2. The linkserver expands to links stored in its database (more functionality can be added in the future).
3. The linkserver database will be build for all bibliographical entries of CIS for which the SICI identifier can be constructed.
4. The linkserver will be programmed in Python.
5. The linkserver runs on VTEX supported server (mirroring in the CIS site can be added in the future).

### 3 Timeline

Step 1. Initial analysis: the agreements on the format of the queries to linkserver, methods to fill the database [2 month].

Step 2. Programming of the linkserver functionality [2 month].

Step 3. Building the linkserver database [2 month].

Step 4. Final launch and final adjustments [1 month].

## 4 Budget

Step 1: 1320\$.

Step 2: 3960\$.

Step 3: 2640\$.

Step 4: 880\$.

Total: 8800\$.

## 5 Future enhancements

The linkserver can be used to add more functionality in the future. The natural extension, now missing in CIS, is to have *publicly accessible identifiers* for all items in CIS. This would allow external links to them from external web resources (like linkservers). Then the authors of the publications could be able to link to the corresponding entries in CIS and so provide links (through CIS and the linkserver) to items related to their publications.

In fact, the identifiers should be visible as active web links next to corresponding bibliographical items, when those are shown in screens exposing search results. Clicking on the identifier in the browser, the page of the bibliographical item could be exposed. This page can show at least the item and links provided by the linkserver, but here is easy to add more activity and more content in the future, like an abstract, reviews, bibserver, etc.

# CIS Proposal for Web Development 2006: Cross Referencing between CIS and Related Fields and Services

Andrew McCallum  
mccallum@cs.umass.edu  
<http://www.cs.umass.edu/~mccallum>

## 1 Summary

We propose to obtain a large collection of bibliographic data from fields neighboring statistics, make it available to CIS and the IMS Biobibs Project in structured form, and perform cross-referencing (name authority) sufficient to integrate it with existing CIS databases. The new data will be obtained by extensive Web spidering for research articles in PDF and Postscript, as well as our existing spidering relationship with the Cornell arXiv.

Whenever we are able to obtain the full-text of these new articles, we will also extract references from the end of the article, extract the separate fields for title, authors, journal, year, pages, etc, and perform reference linking—enabling us to provide lists of articles that cite articles in the CIS or IMS Biobibs.

Furthermore, in a limited number of cases (about half), we will be able to provide homepage URLs for those authors that have them, as well as their email addresses.

We expect to provide about 100,000 papers and 100,000 authors, and potentially much more.

Finally, we will also create links from our own CiteSeer-like computer science research paper search engine into CIS and IMS Biobibs—guiding more Web traffic to [statindex.org](http://statindex.org) and [imstat.org](http://imstat.org).

The work will be done at the University of Massachusetts, however, we will coordinate with VTEX for input of data to the IMS Biobibs Project and/or CIS, and for long term maintenance needs to be determined by IMS/CIS.

## 2 PI Qualifications and the Rexa Project

Associate Professor McCallum directs the Information Extraction and Synthesis Laboratory (IESL) at the University of Massachusetts Department of Computer Science, with 8 PhD students, 2 postdoctoral fellows, and 2 full-time staff software engineers. He is the author of more than 50 research papers on machine learning and natural language processing, and is on the editorial board of the Journal of Machine Learning Research. Prior to his arrival at University of Massachusetts he was Vice President of Research and Development at a 170-person start-up company that used machine learning methods to perform information extraction from the Web.

The project proposed here is enabled by leveraging substantial software infrastructure and bibliometric extraction research being funded by a \$2.6 million NSF ITR project within IESL. The goal of this NSF project is (1) to fund new research in statistical natural language processing targeted at bibliometric data and joint inference across the multiple steps of a language processing and data mining pipeline, and (2) to create an enhanced alternative to CiteSeer and Google Scholar. The Rexa project was begun in 2003; the original NSF ITR goes through 2007; the PI plans to continue the project indefinitely, and has already obtained follow-on supporting grants from NSF and elsewhere.

The new service is called Rexa, and is available at <http://rexa.info>. It currently requires an invitation and password for access. These can be obtained by contacting McCallum.

The core of the research in the NSF project is to intimately integrate information extraction (IE) and data mining—enabling newer, more accurate, more ambitious mining of distributed unstructured textual data sources. Such efforts historically usually have limited success because IE is never perfectly accurate, and data mining is brittle. Our research uses robust probabilistic models, confidence estimation, and joint inference across IE and data mining so that uncertainty and multiple IE hypotheses can be used to improve mining (“bottom-up”), and emerging mined patterns can be used to improve IE (“top-down”). We are making this approach practical in real-world systems—including research in new probabilistic inference and parameter estimation methods along the way.

The core of the application in the NSF project is building an enhanced research paper search engine for computer science—an enhanced alternative to Google Scholar, CiteSeer, ACM Portal, etc.<sup>1</sup> There are many improvements (including better inter-

---

<sup>1</sup>The PI believes that—just as it is important that we have many national newspapers, not just one—it is very important that there be multiple research paper search services. Each has different strengths, weaknesses, biases and emphases. So I see these other services as colleagues, not competitors. For example, the PI has co-authored grants with Paul Ginsparg, and stay in contact with Lee Giles.

face, higher accuracy, tagging, etc) but the main enhancement is that Rexa will know about many more first-class, de-duplicated, cross-referenced object types: not only papers and their citation links, but also people, grants, universities, conferences, journals, research groups, topics. We will then be able to mine this inter-connected web of related objects to understand (a) how ideas travel through social networks of researchers, (b) map out influence and impact of papers, people, departments, grants and research areas, (c) map out sub-fields, and automatically create introductory guides, lexicons, related work, (d) extremely effective people search, (e) automatic paper/grant reviewer suggestion, etc. Papers, people, grants and topics are supported now. Our person coreference system recently won first place in an official competition where other competitors included CMU, Columbia, University of Maryland and Fair Isaac.

Rexa's operation includes:

- Obtaining open access full-text papers by spidering web sites related mostly to computer science, but also including some math, statistics, physics, economics, social science, linguistics, biology, medicine, etc. Our highly efficient spider can download over a million PDFs in just a few days.
- Converting PDF and Postscript to and XML representation for plain text.
- Automatically locating research paper “headers” and “references”.
- On these headers and references, performing information extraction of about 14 different bibliographic fields (author, title, journal, year, etc) using a highly sophisticated machine learning method based on conditional random fields (Lafferty, McCallum, Pereira, 2001, ICML).
- Performing cross-referencing of article citations and author names using sophisticated graph partitioning methods, parametrized edge weights, and parallel processing.
- Depositing all this information into an SQL database and indexing the text fields with Lucene.
- Making the service available through a Web site based on Java, JDBC, Javascript and AJAX.

IESL makes use of over 100 Xeon-class compute servers to perform these tasks.

The currently-available Rexa system has over 7 million research paper references and over 879,000 de-duplicated authors.

### 3 Project Description

The CIS has goals that include: (a) gathering additional bibliographic metadata from new collections, (b) adding value to existing entries, including links to other related articles, (c) providing links to external sources of information related to the entries in the index, for example contact information for the authors. The IMS Biobibs project has the same three needs.

We propose a project that will touch on all three of these needs with some sensible initial steps. We hope that this project will be the beginning of an ongoing supportive and collaborative relationship between Rexa and CIS and the IMS Biobibs project.

The Rexa project aims to be a good community citizen—among both researchers and other providers of bibliographic information services. We believe that research communities benefit when such service providers coordinate and work together, while focusing on the needs of their own constituents. Rexa’s focus is on the computer science literature. Of course, there are long-established and growing links between computer science and statistics that make coordination between CIS/IMS and Rexa interesting.

Our proposed CIS/IMS project consists of the following.

- Spidering additional Web sites suggested by CIS/IMS as having open-access research paper content that would be useful to CIS/IMS. This would likely include non-journal as well as journal content. This will undoubtedly include the arXiv at Cornell, especially its PR and ST sections. We have not yet written the necessary code to spider from the arXiv (it has unusual URLs with a certain kind of re-direction), but we already have a relationship with the arXiv, as well as the necessary permissions and access ports. We could obtain more than 100,000 full-text PDFs.
- Extract bibliographic information from these PDFs—from article headers and references—including title, authors (first, middle, last names), journal, conference, series, volume, number, pages, year, editor, location, publisher and institution. No extraction process (whether by machine or by hand) is perfect. Our conditional-random-field-based extractors are currently the top performers in benchmark tasks tested in the community, obtaining an average of over 97% accuracy.

Of course, for those records obtained from research paper headers (as opposed to references sections), we would also be able to deliver the URL of the full text of the article.

- Spider for BibTeX files as directed by CIS/IMS (whether from BibServer efforts, or from elsewhere), and include their bibliographic records in the subsequent processing.
- Perform research paper reference linking (co-reference, or name authority) between extracted bibliographic records and CIS/IMS bibliographic records. We plan to begin with author records for 900 IMS fellows and 5000 IMS members. This reference linking would enable:
  - the identification of which records are new, and could be added to the CIS,
  - the identification of records already in the CIS, for which we may now know an open-access URL for the article’s full text,
  - the augmentation of existing CIS records with fields that may have been missing from CIS records,
  - lists of research papers that cite papers in the CIS,
  - lists of papers that are cited by papers in the CIS.
  - URLs for many of these papers.

Many of these newly provided links will be between statistics and computer science articles; some will also include papers from statistics, math, social science and other related disciplines.

- Add to Rexa hyperlinks to CIS and the IMS Biobibs Project, so that authorized CIS members can access CIS pages, and all users can access IMS Biobibs pages, directly from relevant Rexa pages. This will drive traffic to [statindex.org](http://statindex.org) and [imstat.org](http://imstat.org) and help users in neighboring fields discover these sites, and encourage new subscriptions to CIS and new membership of IMS.

A significant consequence of CIS/IMS/Rexa collaboration on Biobib data associated with authors who are IMS members should be the provision of name authority in CIS for such authors. Future work could include provision of further name authority e.g. to coauthors of IMS members, to ASA members, and beyond, gathering biographical information on authors (including institutions, electronic and U.S. mail contact information, former institutional affiliations with dates, advisers, etc).

Future work could also include topical analysis to automatically assign new keywords to articles, automatically place them in a topic hierarchy, augment existing topic hierarchies, and perform analysis of how topics are related to each other.



## 4 Deliverables and Timeline

The project will take about 9-12 months.

**Months 1-3** Have dialog with CIS/IMS about choice of open-access web sites to spider. Perform extensive web spidering. Determine which PDFs are research papers and which are not. Report to CIS/IMS on results, obtaining additional and revised list for further spidering. Develop new code to access the arXiv. Spider requested .bib files, and implement methods to incorporate these records into further processing steps.

**Months 4-7** Make improvements to pdf2text, segmentation and extraction code to handle any new formatting issues. Set up and run segmentation and extraction code on CPU server farm. Deposit data into SQL database. Run research paper co-reference linking on server farm, and depositing results into SQL database.

**Months 8-11** Run machine learning methods for author homepage-URL-finding. Devise format for delivery of data to CIS/IMS. Create method and write code to incorporate hyperlinks to CIS/IMS into the Rexa search engine. Deliver resulting data to CIS/IMS. Aid CIS/IMS technical personnel in interpreting and incorporating the data, in coordination with VTEX for long-term maintenance of incorporation process.

## 5 Budget

The project will require significant effort on the part of a research staff programmer. Fortunately the efforts of this person will significantly leverage existing resources from our past research, our existing software, and available computer hardware and storage resources.

We have also obtained leverage in the form of 50-50 matching funds from IMS.

The budget below, for approximately \$10,000, covers just one-half day per week of our staff programmer for 10 months. It includes his salary, standard required UMass benefits, and the standard required departmental IT support fee (system maintenance, email, network bandwidth, backups, etc, per head-count, pro-rated for his percentage time on this project).

Our proposed billing is for \$5,000 from CIS, plus \$5,000 matching funds from IMS.

Programmer salary 8.4% of calendar year	4,543
Fringe benefits	1,574
Computer maintenance	252
Total direct costs	6,369
Indirect costs (57%)	3,630
Total Cost	9,999

Our hope is that this project will form the beginning of relationship involving CIS, IMS and Rexa that will continue in the years ahead, and that there will be further funded opportunities together in the future.

## ECP Volume 10, 2005

The figures for submissions and publications over the last five years for ECP are as follows.

Year	Submissions	Volume	Papers published
2005	71	10	30
2004	67	9	20
2003	51	8	21
2002	34	7	22
2001	34	6	12

The mean time to final acceptance for papers published in 2005 was 176 days.

At the end of 2005 Ted Cox and I retired as editors of EJP and ECP. We wish to thank the Associate Editors and referees for their work in the continued success of these journals.

Martin Barlow  
Editor

## Report EJP

In 2006 the number of submissions increased in the first six months to 87 papers. This increase has not lead to an increase in good papers so the perspective will be that acceptance rates go down and since Associate Editors and referees have a higher load, it will be a difficult task to keep the median time for a decision low.

There seems to be the need for a journal publishing correct and well-written articles which do not have a high level of novelty of the mathematics but might still be useful as reference.

Andreas Greven

EJP Vol 10 2005

82 papers submitted in 2005, up from 71 in 2004  
46 papers published in 2005, up from 28 in 2004

Of papers published in 2005:

mean time to first report: 247 days

median time to first report: 174 days

mean time to final decision: 262 days

median time to final decision: 250 days

Ted Cox

## Report from the JCGS Management Committee for IMS, June 2006

The JCGS Management Committee met at the JSM, August 2005. Here is a summary of the meeting and issues that arose that may need to be addressed at the Joint Meeting of CoP and ASA Editors.

- The current Editor, Luke Tierney, reported on the state of the journal:
  - The last year saw 176 submissions and 51 articles published. The increase of submissions was dramatic, up from 110. The quality is not diminishing which suggests we may want to consider increasing the page limit.
  - Over the past few years there have been approximately 45-50 papers are published each year. The increase in submissions for 2005 of approximately 15 submitted papers per month, produces an acceptance rate of between 25-30%.
  - The backlog of articles is about 9 months, and median time to first review is about 4.5 months.
  - The editor's budget request for the upcoming year will depend on the new editor. For 2005 the budget included \$4000 request from the editor, and this was raised to \$4600 for 2006, plus an additional \$1000 for th editor changeover.
  - Luke is spending about 15-20 hours per week on editorial work, which is a substantial commitment. Ways to reduce the burden may need to be addressed for a new editor to accept the role. Allentrack, or the addition of an editorial assistant might help.
  - There is continued concern about the small number of graphics submissions.
  - There was little change in the editorial board. There are about 35 associate editors.
- One major activity in the past year was to begin the process of finding a new editor (2007-2009) to begin reviewing new articles in July 2006, taking over from Luke Tierney. David Van Dyk (UC-Irvine) has been approved as the new editor by ASA, IMS and Interface. As part of the negotiations with David the Management Committee agreed to provide an editorial assistant and to switch to an electronic review management system.
- The budget was discussed. A major part of the next year's budget is likely to be to cover the costs of switching to AllenTrack. Whether the switch happens depends on the new editor. The cost of color also was discussed. To be a successful graphics journal more color is needed. A low cost color solution may be worth investigating. Total revenue for 2005 was \$106,500 and expenses \$103,000, yielding a net profit of \$3,500.
- There was a discussion on continuing to raise the profile of JCGS, in relation to graphical and computational articles.
- Double-blinding of ASA articles made the discussion again this year. The general opinion is that double-blinding is increasingly unrealistic today. Authors' identities can often be easily found using the Internet. Double-blinding is the policy of the ASA. Although it might be difficult to strictly enforce it does convey a message to reviewers that author's identities should not be a consideration in their review.
- The relationship between JSS and JCGS was discussed. ASA has been considering making JSS an independent journal. The board approved this in the latter part of 2005.
- JCGS received a new cover last year, one that prominently displays the logos of the three sponsors.

## Probability Surveys Annual Report 2006

Volume 2 (2005) contained 13 papers (549 pages), an increase over Volume 1 (2004) which contained 6 papers (392 pages). As of May 12 2006, Volume 3 contains 6 papers (229 pages) and 7 more papers are under review. This seems to be an acceptable albeit slow rate of growth.

As well as traditional survey papers, we are soliciting write-ups of research summer school lecture notes, and in 2005 published one such (by Le Gall).

Acceptance percentages are not comparable to those of other IMS journals for various reasons. Of papers formally submitted in 2005, there were 4 papers rejected as "more like new research than a survey paper". Such rejections are generally made very quickly, directly by Editor or after consultation with an Associate Editor. I also receive (and welcome) informal enquiries about suitability of draft papers, which obviates later need for rejection of formal submissions.

David Aldous