Establishment of a Graduate Interdisciplinary Program in Statistics

I. Program Name, Description and CIP Code

A. Degrees, Department and College and CIP Code

CIP Code: 270502

The proposed program is named Statistics Graduate Interdisciplinary Program (Statistics GIDP). Being interdisciplinary in nature, it will serve all colleges and academic units within the UA. It will offer M.S. and Ph.D. degrees in Statistics, as well as a Ph.D. Minor in Statistics.

B. Purpose and Nature of Program

The proposed Graduate Interdisciplinary Program (GIDP) in Statistics will fill a vital need at the University of Arizona and for the State of Arizona in general. There is currently no functioning M.S. or Ph.D. program in statistics at the University of Arizona – a highly anomalous situation among Research I Universities, especially among land-grant universities with medical schools. The University’s unique institutional structure of interdisciplinary programs provides a natural and resource-efficient mechanism to train students in statistics, and to satisfy the ever growing need for statistical data analysis. This is now especially important in the biological and biomedical sciences where many research groups at the University are hampered by the lack of collaborative support from professional statisticians and appropriately trained students. Such collaborations are often a requirement of research grants from the NIH and other funding agencies, and the current situation in which in which these services have to be subcontracted to outside universities is a serious loss to University funding and prestige.

The proposed GIDP will provide the necessary training for M.S. and Ph.D. students in Statistics, a Ph.D. Minor in Statistics, and also the opportunity for supplementary Masters training for Ph.D. students in the life sciences and other graduate programs. Although approximately fifty graduate courses may be found among the listings of over a dozen departments in many different Colleges, there is a paucity of trained statisticians to teach all of them, and these departments have placed little emphasis on recruiting graduate students interested in statistics in recent years. The GIDP will provide the appropriate venue to rationalize and utilize these courses and, at the same time, provide a focus of interest for current faculty and a means of attracting new, statistically oriented, faculty to the University. For example, in the 2004 grant year, faculty from Bio5 subcontracted in excess of one million dollars of grants for outside statistical support.
Over the last several years, faculty in the Departments of Mathematics, Ecology and Evolutionary Biology, Economics, Epidemiology, and Systems Engineering (and other departments as well) began to examine the campus-wide statistics situation. They established an interdisciplinary Statistics Colloquium that has attracted speakers and attendees from across the campus, strengthened ties among the campus-wide statistics community, and has provided a mechanism for obtaining advice from nationally prominent statisticians. The overwhelming opinion of the participating faculty is that Statistics at the University of Arizona should be organized as a separate, interdisciplinary, academic unit, with the clear label “Statistics” attached to it, and with a core of theoreticians and a flexible network of statisticians and statistics users and consultants.

C. PROGRAM REQUIREMENTS

The M.S. Program in Statistics

Since the basic graduate level (500 level) courses are already offered by a variety of departments — especially departments in the Colleges of Science, Public Health, Engineering, and Agriculture and Life Sciences — it will be possible to immediately advertise an interdisciplinary Master’s program in Statistics at U/A. This degree formally still exists, and can be awarded by the GIDP. Here are the basic requirements for this degree:

(a) A Bachelor’s Degree, either in a mathematical field or a field that makes significant use of quantitative methods.

(b) At least two semesters of Calculus (at the level of Calculus-I, II, i.e., MATH 125, 129), and one semester of Linear Algebra (at the level of MATH 215).

(c) Some exposure to elementary statistics, at least at the level of MATH 263. A semester of upper-division probability is strongly recommended, but could be made up as a deficiency, by taking MATH 464/564.

The M.S. Degree in Statistics Requirements:

A minimum of 33 units (36, if the Master’s thesis option is chosen). These units will be made up as follows.

Core M.S. Courses:
Theory of Statistics (MATH 566 or AREC 517).
Regression and Multivariate Analysis (MATH 561 or EPID 684).
Experimental Design (MATH 571 or ANS 553 or PSYC 507C).
Statistical Consulting (MATH 572).

Additional Courses Required:
(i) Three courses chosen from quantitative methods courses offered by other departments.
(ii) A six-unit Master’s thesis, or a three-unit smaller-scale project that involves a significant write-up and a final oral report.
(iii) Three courses from the following:

2
Time Series (MATH 562 or SIE 532).
Categorical Data Analysis (MATH 570)
Applied Stochastic Processes (MATH 568 or SIE 520 or SIE 531 or SIE 531)
Advanced Topics in Biostatistics (EPID 576B-C or EPID 676 or ECOL 581)
Data Management and the SAS Programming Language (EPID 576D)

All the listed courses already exist at UA.

As part of the core requirements, the students need to pass an exam on the use of current statistical software packages.

The Ph.D. Degree in Statistics Requirements:

To be admitted into the Ph.D. candidacy, a student must pass a written comprehensive Ph.D. Qualifying Examination, as determined by the faculty each year. In addition to the core courses required for the Master's program, the Ph.D. students will be required to take the following courses:

(i) Probability Theory (MATH 563A).
(ii) Theoretical Statistics (MATH 567A)
(iii) Theoretical Statistics (MATH 567B)

and one of the following courses:

(iv) Probability Theory (MATH 563B)
(v) General Linear and Mixed Effects Models (EPID 684)
(vii) Advanced Topics in Biostatistics (EPID 676)

Students will be required to take at least 9 additional semester hours of approved statistics courses. Furthermore, the students need to take 18 additional hours of dissertation credit and other examination(s) required by the Graduate College, as well as the requirements for a Ph.D. minor as stipulated by the Minor department.

D. CURRENT COURSES AND EXISTING PROGRAMS

Current Courses. Currently, as listed in Appendix 1, there are over 50 graduate-level statistic courses offered through 13 departments representing five colleges, Agriculture and Life Sciences (Agricultural and Resource Economics, Animal Sciences, Entomology, Renewable Natural Resources), Engineering (Hydrology and Water Resources, Systems and Industrial Engineering), Public Health (Epidemiology and Biostatistics), Social and Behavioral Sciences (Economics, Psychology, Sociology), and Science (Ecology and Evolutionary Biology, Geosciences, Mathematics).

The creation of a coherent training program and coordinated group of faculty, as embodied by a, GIDP, will provide the ideal mechanism to coordinate and rationalize this
apparent wealth of courses. This will enable students to make the more informed and
effective choices of course-work necessary to obtain a comprehensive training in a
particular approaches and methodologies in statistics.

Existing Programs. A Ph.D. degree in Statistics and an M.S. degree in Statistics still
exist in the books. However, they have not been awarded for quite some time.

E. NEW COURSES NEEDED

No new courses are needed to initiate the Statistics GIDP. In time, as the GIDP grows,
needs for additional courses may arise. In many cases, such courses may be offered as
Special Topics Courses through a number of departments, as has been the case in the
past.

G. REQUIREMENTS FOR ACCREDITATION

Not Applicable

II STUDENT LEARNING OUTCOMES AND ASSESSMENT

Intended outcomes and plan to achieve these:
The MS in statistics will assume basic competency in the theoretical foundations
of mathematical statistics, in particular distribution theory, estimation,
hypothesis testing, and model-fitting. Successful MS students will also have
competency with using the standard tools of statistical analysis such as the general linear
model, nonparametric data analysis, model selection, and statistical computing. Such
training will be achieved through a basic core set of courses coupled with supervised
research with a faculty mentor (or mentors).

A second critical training component is that, by the very nature of their occupations,
statisticians must be highly interactive with non-statisticians. Thus, students must be able
to interact with an extremely wide audience of potential users, from academic researchers
in very diverse departments, to fellow graduate students from other departments seeking
advice, government and business officials, and the general public. Students will thus be
expected to develop strong communication skills, in particular, those dealing with
assessment and formulation of posed problems, strengths and weaknesses of various
potential approaches for the analysis of a particular data set, and presentation and
interpretation of the results. A key component to this training will be very active
statistical consulting, initially in conjunction with Statistics faculty. A second component
of this training will be a weekly seminar in which students learn to present their analysis
to faculty and fellow students.

The Ph.D. in statistics will assume all of the required competencies of the MS

4
Degree, plus a much deeper level of the theoretical underpinnings of statistics and its broad applications. In particular, Ph.D. students will be expected to be capable of developing new statistical methods for the analysis of complex data, rather than simply applying existing approaches, and to assess the strengths and weaknesses of such newly-developed approaches.

Assessment:

Incoming students will take an entering diagnostic exam to provide a benchmark for the knowledge of entering students as well as to help them assess the current level of their skills. Both MS and Ph.D. students will take a written qualifying exam following completion of the core curriculum, which will assess competence in the core concepts. Ph.D. students will also take an oral exam testing both competence in advanced topics as well as in specific areas of their field of interest. Yearly meeting of the statistic core faculty will be held to assess student performance on these exams, with critical feedback provided to the graduate and curriculum committee for action, when appropriate.

One, three, and five-year follow-up assessments will be conducted among our graduates. These will be web-based, but with active contact from the statistic staff to ensure 90% return on the one-year and upwards of 80% on the three- and five-year responses. These surveys will allow students to self-assess the performance of the Statistics GIPD in training them for their professional careers. Feedback from these surveys will be forwarded to the graduate and curriculum committees for appropriate action.

IV. STATE'S NEED FOR THE PROGRAM

The State of Arizona and its three Universities have targeted the biomedical sciences and biotechnology as vital areas for economic growth. The University of Arizona and Arizona State University are currently developing major, institute-based initiatives in these areas, which are complemented by the State supported TGEN enterprise. The strategic and economic importance of biomedical science and technology, and the need for the appropriately trained workforce, has been clearly identified in the Flinn-Batelle Roadmap. Central to advances in the biomedical sciences and the development of new drugs is biological data analysis, and the proposed program will provide much needed personnel well trained in these techniques. ASU and NAL currently teach statistics through their Departments of Mathematics and Statistics, and the proposed graduate interdisciplinary program at the University of Arizona will provide a natural complement to these units. The paucity of well-trained statisticians is not unique to the State of Arizona, and enhancing the State's workforce in statistics will be of enormous benefit to all of the Universities, and can help attract biomedical and biotechnology companies to the region.

Anticipated Student Enrollment
### FIVE YEAR PROJECTED ANNUAL ENROLLMENT

<table>
<thead>
<tr>
<th>Year</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS students</td>
<td>5</td>
<td>10</td>
<td>20</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>Ph.D. Students</td>
<td>6</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>20</td>
</tr>
</tbody>
</table>

We anticipate that a number of Ph.D. students in the life sciences (Colleges of Public Health, Science, Agriculture) will also seek a masters degree in statistics. This is fairly common at our Peer institutions especially for life science degrees that are highly quantitative (e.g., Ecology, Genetics, wildlife management, bioinformatics). We expect the master to be up and running fairly quickly, while some inevitable delay is expected for Ph.D. level students. In particular, we expect it to take at least a year after the initiation of the program for suitable advertisements and recruitment activities to attract the quality of students we are striving for, hence the delay in projected Ph.D. enrollment until year two.

The above projected numbers do not include students enrolled in the Ph.D. minor in statistics. We expect these numbers to roughly be a quarter to half that for the major, namely 5-10 Ph.D. minors in years 4 and 5.

**What are the local, region and national need for this program? Provide evidence of the need of the program. Provide an assessment of employment opportunities for graduates of the program during the next three years.**

The local, regional, and national needs for this program are critical. Perhaps the best indicator for the need (locally and regionally) for statisticians is that in 2004, over one million dollars in grant funds (in just the area of genomics) awarded to UA researchers had to be subcontracted to other institutions for the statistical analysis. This is also a national problem, as most major universities are scrambling to implement bioinformatics programs and are facing a critical shortage of appropriately-trained individuals, esp. those with the appropriate statistical expertise. Formally trained statisticians have always been in short supply, and is high demand. Employment opportunities are thus exceptional for graduates of the program.

### IV. APPROPRIateness FOR THE UNIVERSITY

#### A. EXISTING PROGRAMS AT OTHER CAMPUSSEs

**EXISTING PROGRAMS IN ARIZONA:** At ASU the Mathematics Department is now called the Department of Mathematics and Statistics, with a number of statisticians within
the department. The same is true of NAU. There is no separate interdisciplinary graduate level statistics program at any of the major academic institutions in Arizona.

B. PROGRAMS OFFERED IN OTHER WICHE STATES

ALASKA: Has no separate statistics program in any of its major academic institutions.

CALIFORNIA: There are fine statistics programs in a number of major universities in California. UC Berkeley, and Stanford University have the two top Statistics Departments in the nation. Many of the breakthroughs in theoretical and applied statistics have come from these two institutions. UC Davis has a Department of Statistics, as well as a Program in Biostatistics, offering M.S. and Ph.D. degrees in both Statistics and Biostatistics. UCLA has a Department of Statistics and a Department of Biostatistics, both of fine reputation. UC, Riverside has a Department of Statistics.

COLORADO: Of the two major schools, Colorado State University has had a good Department of Statistics for over thirty years, but University of Colorado at Boulder does not have separate Statistics Department.

HAWAII: None of the campuses of the University of Hawaii has a separate Department of Statistics.

IDAHO: Does not have any Statistics Department among its universities.

MONTANA: No Department of Statistics among the universities in this state.

NEVADA: No Department of Statistics among the universities in this state.

NEW MEXICO: The main state university is State University of New Mexico, which has a Department of Mathematical Sciences, but no separate Department of Statistics.

NORTH DAKOTA: Among the Universities in the State, only North Dakota State University at Fargo has a Department of Statistics, with a small faculty.

OREGON: Of the two major universities in the state, Oregon State University (Corvallis) has a Department of Statistics, while University of Oregon (Eugene) does not.

SOUTH DAKOTA: None of the institutions has a Department of Statistics.

UTAH: Neither University of Utah (Salt Lake City) nor Utah State University (Logan) has a Department of Statistics.

WASHINGTON: University of Washington (Seattle) has two highly reputed graduate statistics programs—a Department of Statistics, and a Department of Biostatistics. The Department of Biostatistics has been in place nearly thirty years, while the Department of
Statistics was established a few years later. Washington State University (Pullman) also has a Department of Statistics, which awards a Master’s Degree, but not a Ph.D.

WYOMING: University of Wyoming (Laramie) has a Department of Statistics, which offers a Master’s Degree and a Ph.D.

Almost all the departments of statistics listed above run Statistical Consulting Services.

EXPECTED FACULTY AND RESOURCE REQUIREMENTS

FACULTY

Given the large number of faculty with some interest in statistics at the University of Arizona, we plan to model this proposed GIDP after the highly successful Applied Mathematics GIDP. Applied Math has a set of core (or full) members and a set of associate members. Full members are those involved in teaching core courses, mentoring students in the program, and others who have agreed to be significantly active in the program. Associate members are those with a general interest in statistical issues who wish to be fully informed of program’s activities, and who may choose to rotate in as full members at some point in the future. Likewise, full members may choose at rotation out to associate members. This structure has proven highly successful in Applied Mathematics and allows faculty to control their level of commitment while still providing for both a general community of statisticians and a core group of highly activity members.

Appendix 2 lists 16 faculty (from 11 departments covering 4 colleges) who will initially form the core faculty.

Additional faculty – describe needed during the next three years and the schedule for addition of these faculty

There is an aggressive plan for hiring additional statisticians on the UA campus (through the BIOS Institute and the Colleges of Science, Engineering, Agriculture and Life Sciences, and Public Health). We fully anticipate these new hires also becoming members of the Statistics GIDP. Indeed the existence of this proposed GIDP will be a key in the successful recruitment of outstanding individuals.

Current FTE students and Faculty – in department unit in which program will be offered

NA

Projected FTE students and Faculty (next three years) in department unit program offered.

NA
LIBRARY

Holdings relevant to statistics are found in the University of Arizona Science, Main, and Health Sciences libraries. These libraries have complete series of the major statistics journals (Appendix 3), as well as access to back issues of most of these journals electronically under JSTOR. No additional holdings are requested nor needed.

PHYSICAL FACILITIES AND EQUIPMENT

Existing Physical facilities. Equipment, special classrooms, laboratories, computer facilities.

No special classrooms or laboratories are required. On-campus computer resources and site-licensed software are sufficient for the majority of needs of the program.

Additional facilities required or anticipated.

While potential future hires may require additional computation power (such as a small cluster of linked computer nodes), this can be easily addressed in their start-up packages. Indeed, recent mathematical hires in Ecology and Evolutionary Biology with similar computer needs were easily accommodated within standard start-up funding.

OTHER SUPPORT

Other support now available: Support staff, University and non-university assistance.

Core faculty (Appendix 2) are in place on campus.

Other support needed, next three years. List additional staff needed.

A half-time administrative assistance/program coordinator is needed.

FINANCING

The costs of this proposed GIDP in statistics are as follows:

<table>
<thead>
<tr>
<th>Description</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Half-time program coordinator</td>
<td>25,000</td>
</tr>
<tr>
<td>Administrative Supplement for GIDP Director</td>
<td>7,500</td>
</tr>
<tr>
<td>Departmental Seminar series</td>
<td>5,000</td>
</tr>
<tr>
<td>Misc. supplies, operations</td>
<td>5,000</td>
</tr>
<tr>
<td>Student recruitment</td>
<td>5,500</td>
</tr>
</tbody>
</table>
SUPPORTING FUNDS FROM OUTSIDE SOURCES

The Deans from the College of Science (COS), College of Agriculture and Life Sciences (CALS), the College of Public Health (COPH), and the College of Pharmacy have each pledged $8,500 per year. COPH will re-evaluate their contribution after year two, which is certainly reasonable and prudent. The College of Engineering has pledged $4,000 for three years to support a 1/4 time TA. The VP for research has agreed to help us cover any remaining dollar differences. Likewise, BIOS has been extremely supportive, pledging roughly $50,000 for FY'96 for student support in a biostatistics training initiative.

OTHER RELEVANT INFORMATION

Space: The Dean of the College of Medicine has promised space for two offices and adjacent space to house 8 graduate students.

Graduate student support. We fully expect the majority of our students to be supported as RAs on various grants, with the remaining likely supported by being more conventional TAs for existing statistics courses. Additional support is very likely forthcoming. For example, BIOS alone has contributed roughly $50 K for FY'96 for student training in biostatistics.

ADMINISTRATIVE STRUCTURE

Executive Committee

The proposed GIPD will be run by an executive committee. The initial makeup of our committee includes board coverage of University Colleges and is as follows:

Dr. Bruce Walsh (Chair)
Depts. of Ecology and Evolutionary Biology (COS), Molecular & Cellular Biology (COS), Plant Sciences (CALS), Animal Sciences (CALS), Epidemiology and Biostatistics (COPH). Associate Department Head, EEB. Member, Genetics, Applied Mathematics, and Insect Sciences GIDP. BCO 3

Dr. Ron Askia
Head, Department of Systems and Industrial Engineering (ENG)

Dr. Rabindra N Bhattacharya
Dept of Mathematics (COS)

Dr. Sylvan Green
Linda McCartney Breast Cancer Chair in Biometry Cancer Center (MED), Dept. of Epidemiology and Biostatistics (COPH).
Dr. Alan Ker
Head, Dep. of Agricultural and Resource Economics (CALS)

Dr. Duane Sherrill
Epidemiology and Biostatistics (COHP).

Dr. Michael Tabor
Chair, Applied Mathematics, Dept. of Mathematics (COS), Dept of Physics (COS), Member BIOS

Dr. Arthur (Larry) Wright
Dept of Mathematics (COS)

Program Bylaws

The proposed bylaws for the Statistics GIDP, which are modeled after those for the GIDP in Cancer Biology, are given in Appendix 4.