NEW UNDERGRADUATE PROGRAM PROPOSAL

ILLINOIS INSTITUTE OF TECHNOLOGY

The following information is required by the Undergraduate Studies Committee to approve new programs. After approval by UGSC this form should be routed to Faculty Council for approval and then the Provost's office.

College(s): College of Science

Department(s): Applied Mathematics

Date: 02/12/2018

Approvals Required

(1) Academic Unit Head(s):

(2) Dean(s):

(3) Undergraduate Studies Chair:

GENERAL INFORMATION

Program Title: Bachelor of Science in Statistics

Program Scheduling: *In what semester will students start to be admitted?* Fall 2018

Total Program Credit Hours: *126 hours minimum* 126

Program Description: *Provide a brief narrative of the program content (use as much space as needed).*

The Applied Mathematics Department proposes a new Bachelor of Science degree in Statistics. We believe that the proposed B.S. in Statistics will address an unmet demand from prospective students, and will consequently increase the number of undergraduate students in our programs. Currently, we do not offer a technically-intensive, mathematically-rigorous undergraduate degree designed to attract prospective students interested in careers in Statistics or Data Analytics. With the proposed new degree in place, we can produce competent graduates who can satisfy the significant and increasing demand in the job market for talent in Statistics, Analytics, and Data Science. This degree will also serve to better prepare undergraduate students who wish to pursue graduate study in these fields.

We intend this program to have a strong computational emphasis, with its graduates taking several programming courses, and its Statistics courses emphasizing statistics computational skills.

The Applied Mathematics Department already offers, in addition to a wide selection of undergraduate mathematical courses, several Statistics courses. For example, we offer, yearly, MATH 475 (Probability), MATH 476 (Statistics), and Math 484 (Regression and Forecasting), which cover the basic Probability, Statistics and Data Analysis methods. Combined with the necessary mathematics courses that our department already offers, we are ready to open this new program.

The Department has voted unanimously in favor of this proposal.

Program Purpose: *Provide details on the intent of the program and its relation to other programs.*

Motivation and Overview

The proposed B.S. in Statistics program is designed for students wishing to obtain enough training in both Mathematics and Statistics, so as to pursue career in Statistics, Analytics, or Data Science in industry, or continue higher education in graduate schools. The program requirements are comparable to the B.S. degrees in Statistics offered within other Mathematics departments.

Given that incoming students at IIT typically have had advanced courses during their high school, a faster and, if preferred, research-oriented program can also be opted for. For example, Camras scholars are typically keen on continuing to graduate studies and participate in research.

Students will be advised in their choice of Statistics electives, according to their future plans: for example, students who are interested in pursuing graduate studies, will be advised to elect Math 400, Analysis.

A sample four-year plan is provided in this proposal.

Admission to the Program

Because the program is designed for high-quality students who are determined to obtain a Mathematics-based degree with Computational emphasis, the Mathematics/Statistics and Computer Science loads are heavy in most semesters. The students must meet the same admission requirements of the other B.S. degree program of our department.

Program Benefits: *State the impact of the program for students and for IIT.*

This program will increase enrollment in mathematical sciences, and will provide students with an interest in statistics, and/or an inclination towards statistics and data science jobs with an explicitly marketable degree with which to pursue those career. With statisticians and data scientists in high demand, with low unemployment and high median pay, in 80K range, students who graduate from this program will be well equipped to pursue successful careers, and will be happy future alumni.

Classification of Instructional Programs (CIP) Code: 27.0503

Required to make the program US Financial Aid Eligible - The CIP code takes the following structure: xx.xxxx Where each x is a number between 0 and 9. This 6-digit code identifies, to the greatest specificity possible, an entire instructional program. The classification scheme seeks to comprehensively address all areas of study. Because of the dynamic nature of education, however, new CIP codes are frequently added to the list. The first 2-digits are the first cut off of detail and describe the general discipline of the program. For example, any program with a CIP that starts with 14 is within the Engineering discipline; anything with a 22 is within the legal discipline. The next 2 digits increase the level of detail, and the final 2-digits provide the highest level of detail.
Find CIP codes at http://nces.ed.gov/ipeds/cipcode

PROGRAM VIABILITY

Competitive Programs: *Indicate other similar programs locally and nationally detail their success.*
Similar Statistics programs are offered by many universities, including (locally) Loyola University and University of Illinois at Chicago.

Market Analysis for Recruiting Students: *Detail what work has been done with UG Admissions to identify and recruit potential students.*
The target market is high school students with mathematical ability and interest in high-paying, high-satisfaction careers, including data science. We will also work with UG admissions to identify and recruit potential students. The current high level of awareness of statistics and data science provides a rich market environment for recruiting students.

Market Analysis for Graduates: *Detail what work has been done with the Career Management Center to identify potential employment opportunities for graduates*.

Statisticians are highly sought after, and such jobs are highly paid, and in high demand. One need only look at sources such as the US News & World Report, where Statisticians are listed as the #1 Best Business Job and the #6 Best Job for 2018. We will, of course, work with the Career Management Center to identify specific potential employment opportunities also.

ACADEMIC INFORMATION

Enrollment Estimates: *Are there enrollment estimates for this program, and if so, what are they and what are they based on? What is the minimum number of students necessary in the program to make the program viable (i.e.to offer classes unique to the program often enough)?*

We do not have enrollment estimates currently, but the new courses the program requires are electives for the Bachelor of Science in Applied Mathematics and/or will be relevant to students of the other programs in the department, so we do not foresee difficulties in this. The program is viable with any number of students; an enrollment of 20 would make it comfortably stable.

Advising Strategy: *Since quality advising is a key component of good retention, graduation and career placement, how will students be advised and mentored? Specifically for interdisciplinary programs, how will advising responsibilities be shared? What student professional organizations will be formed? How will the department work with the Career Management Center to develop industry connections?*

Students will be have a designated faculty advisor, initially Despina Stasi, who will advise them on their choice of courses, their progress in the program, and the selective statistics and applied mathematics electives, according to their future plans. For example, students who are interested in pursuing graduate studies will be advised to take Math 400 Analysis.

The department will encourage and facilitate the establishment of a student chapter of the American Statistical Society (ASA), which can be done at no cost. The students would also benefit from the current student organizations associated with the department, the Society for Industrial and Applied Mathematics (SIAM) and the Association for Women in Mathematics (AWM).

AMAT currently works with CDC to host a once-per-semester alumni career panel, which connects the students with alumni in companies that also hire statisticians. We also have faculty members with personal connections to industry, and to statistics jobs. And, of course, the department will happily work with CMC to develop further industry connections related to this new program.

Course Requirements: *Detail the courses needed for the program including courses currently offered, new courses to be developed (including syllabi), and dependence on courses from other academic units with their commitments to provide these courses on a long-range basis. Include descriptions of laboratories that will need to be developed along with equipment and facilities requirements.*

Required Courses	Credit Hours
Applied Mathematics Requirements MATH 100, 151, 152, 230, 251, 332, 350, 435	26
Statistics Requirements MATH 225, 475, 476, 484, 4XX (time series/forecasting)	18
Applied Mathematics/Statistics Electives MATH 252, 380, 400, 481, 483, CS 422, or other MATH elective.[1]	15
Humanities and Social Sciences Requirements See IIT Core Curriculum, sections B and C (p. 40)	21

Minor Subject Requirement Five related courses from departments outside of applied mathematics	15
Interprofessional Projects (IPRO) See IIT Core Curriculum, section E (p. 41)	6
Computer Science Requirement (CS 115 & CS 116) or (CS 104 & CS 201) or (CS 105 & CS 201), CS 331	7 (or 9)
Natural Science or Engineering Requirement See IIT Core Curriculum, section D (p. 40)	10
Free Electives	8
Total Minimum Hours	126

[1] Applied Mathematics/Statistics Electives are to be chosen after consultation with an academic adviser. Student goals, interests, and course availability should be determining factors in this selection process. Students can take CS 422 Data Mining to replace one Applied Mathematics/Statistics Elective. CS 422 must be taken after CS 331, which is a required CS course in this curriculum. The following courses do not count towards the requirements for this degree: MATH 119, 122, 130, 148,180, 333, 374, 425, 426, 474.

Sample Curriculum/Program Requirements: *Provide a sample semester by semester curriculum and the program requirements, as they would appear in the IIT Undergraduate Programs bulletin.*

Semester 1	**Credits**
MATH 100 Introduction to the Profession	3
MATH 151 Calculus I	5
Computer Science Course[1]	2
Social Sciences Elective	3
Humanities 200-level Elective	3
Total Hours	**16**

Semester 2	**Credits**
MATH 152 Calculus II	5
MATH 230 Introduction to Discrete Math	3
MATH 225 Introductory Statistics	3
Computer Science Course[1]	2
Science/Engineering Course	4
Total Hours	**17**

Semester 3	**Credits**
MATH 251 Multivariate and Vector Calculus	4
MATH 332 Elementary Linear Algebra	3
CS 331 Data Structures and Algorithms	3
Minor Elective 1	3
Social Sciences Elective (300+)	3
Total Hours	**16**

Semester 4	**Credits**
MATH 435 Linear Optimization	3
Applied Mathematics/Statistics[2] Elective 1	3
Minor Elective 2	3
Science/Engineering Course	3

Social Sciences Elective(300+)	3
Total Hours	**15**

Semester 5	**Credits**
MATH 475 Probability	3
Applied Mathematics/Statistics[2] Elective 2	3
Science Engineering Course	3
Minor Elective 3	3
Humanities Elective (300+)	3
Free Elective 1	2
Total Hours	**17**

Semester 6	**Credits**
MATH 350 Intro to Computational Math	3
MATH 476 Statistics	3
Applied Mathematics/Statistics[2] Elective 3	3
Humanities or Social Sciences Elective	3
IPRO 397	3
Total Hours	**15**

Semester 7	**Credits**
MATH 484 Regression	3
Applied Mathematics/Statistics[2] Elective 4	3
Minor Elective 4	3
Humanities Elective (300+)	3
IPRO 497	3
Total Hours	**15**

Semester 8	**Credits**
Math 4XX Time Series and Forecasting	3
Applied Mathematics/Statistics[2] Elective 5	3
Minor Elective 5	3
Free Elective 2	3
Free Elective 3	3
Total Hours	**15**

[1] Students must complete one of the following computer science sequences: CS 115 and CS 116, CS 104 and CS 201, or CS 105 and CS 201.
[2] Applied mathematics/Statistics Electives are to be chosen after consultation with an academic adviser. Student goals, interests, and course availability should be determining factors in this selection process.

Extra Resources Needed

The only necessary extra resources will be provided by the AMAT department: this will not be difficult to achieve as we have recently been adding faculty with the capability of teaching Statistics courses, with this program in mind. Syllabi for Math 225, Math 4XX and revised syllabus for Math 484 are attached. Some specifics follow.

- Currently MATH 483 Design and Analysis of Experiments is offered every four years. To meet the minimal need of the new program, it needs to be taught every other year.
- MATH 225 is a calculus-free statistics course. Its purpose is to introduce the basic statistics concepts to the statistics major early in their program, before they have the calculus background required to take Math 476. This course will be offered yearly.

- Currently MATH 484 covers topics in Regression, as well as Time Series/Forecasting. We propose to split that course into MATH 484, now covering Regression in a more comprehensive manner, as well as some Design of Experiments topics, and propose to add a new course MATH 4XX to cover Time series/Forecasting, which will be a Statistics elective for this program. We expect that the new course will serve undergraduates from our B.S. in Applied Mathematics degree as well as the B.S. in Statistics students.

Program Outcomes and Assessment Process: *Provide the program learning goals and assessment plan (for more information contact the Assessment Office within Academic Affairs).*
Also see https://sites.google.com/a/iit.edu/student-learning-assessment/

The core courses of the program each have their own learning goals and objectives, and in totality these give a general rubric for the program. One metric for program outcomes will be the number of students in the program and another will be their graduation rate (within 4,5,6 years).

Students will:

1-Understand the fundamentals of probability theory, statistical reasoning, and inferential methods

2-Learn how to work with data, describe it, model it, analyze it, and present it.

3-Propose statistical models, evaluate their validity, use them to draw conclusions, and understand their limitations.

4-Have strong computational skills and be able to effectively use statistical computing to answer real-world questions.

5-Learn to communicate statistical analyses and conclusions clearly and effectively.

6-Understand the broader and dynamical societal context in which statistics is used and demonstrate awareness in ethical issues associated with sound statistical practice.

Math 225 - Introductory Statistics

Course Description: An introduction to statistics; data collection, description, visualization and analysis; basic probability; statistical reasoning and inference including hypothesis tests and confidence intervals: t-tests, chi-squared tests, ANOVA, correlation and regression. (3-0-3) (C)

Enrollment: Required for Statistics Major and Minor. Elective for other majors.

Textbooks: Statistics: Unlocking the power of data 2nd edition with WileyPLUS access, Lock et al.

Other Required Materials: JMP or other statistical software.

Prerequisites: None

Objective:

1. Students will become critical consumers of statistically-based results in the media and be able to recognize whether reported results follow from the studies and analysis reported.
2. Students will understand and be able to employ the key concepts of statistical inference: estimation with intervals and testing for significance.
3. Students will able to use statistical software for data analysis, and be able to interpret and draw conclusions from the output.
4. Students will understand basic ideas of formal probability theory and be able to compute probabilities of events for simple examples.
5. Students will be able to demonstrate awareness in ethical issues associated with sound statistical practice.

Lecture Schedule: Two 75-min sessions per week.

Course Outline	**Hours**
1. Data Collection: structure, sampling, experiments and observational studies.	3
2. Data Summary: categorical variables, measures of variability, visualization, z-scores.	3
3. Basics of Theory of Probability: probability laws, random variables, Bayes Rule.	3
4. Confidence intervals: sampling distributions, interval estimates, bootstrapping.	6
5. Hypothesis tests: statistical significance, randomization distributions, confidence intervals	7
6. Normal Distributions	4
7. Inference for means and proportions: tests and confidence intervals for proportions, t-tests.	5
8. Inference for multiple parameters: Chi Square Tests	3
9. Inference for multiple parameters: ANOVA	3
10. Inference for multiple parameters: Correlation and Regression	3

Assessment

Homework/Projects/Labs	30 - 40%
Mid-Exam(s)	30 - 40%
Final Exam	30 - 40%

Syllabus Prepared By: Despina Stasi and Arthur Lubin, 02/02/2018

MATH 484/564 Regression/Applied Statistics

Course Description: This course introduces the basic statistical regression model and design of experiments concepts. Topics include simple linear regression, multiple linear regression, least square estimates of parameters; hypothesis testing and confidence intervals in linear regression, testing of models, data analysis and appropriateness of models, generalized linear models, design and analysis of single-factor experiments.

Enrollment: Required for B.S. in Statistics; Elective B.S. in Applied Math, M.S. in Applied Math, Ph.D. in Applied Math, and all the professional master programs of the AMATH.

Textbooks: M. Kutner, C.J. Nachtsheim, J. Neter, W. Li, *Applied Linear Statistical Models*, 5th edition (with Student CD). McGraw-Hill/Irwin

References:

1. W. N. Venables and B. D. Ripley (2002). *Modern Applied Statistics with S*, 4th Edition. Springer.

Software: R

Prerequisites: MATH 474 or MATH 475 and 476, min grade of C.

Objective:

1. Students will learn about the regression models and basic design of experiments concepts and their applications in various fields of science and engineering.
2. Students will be able to formulate real life problems using regression models.
3. Students will be able to use statistical software to estimate the models from real data, and draw conclusions and develop solutions from the estimated models.
4. Students will learn to use visual and numerical diagnostics to assess the soundness of their models.
5. Students will learn to communicate the statistical analyses of substantial data sets through explanatory text, tables and graphs.
6. Students will learn to combine and adapt different statistical models to analyze larger and more complex data.

Lecture Schedule: Two 75-min sessions per week.

Course Outlines

1. Simple LInear Regression (8 hours)
    - Linear regression with one predictor variable
    - Inferences in regression and correlation analysis
    - Diagnostics and remedial measures

- Simultaneous inference in regression analysis
- Matrix approach to simple linear regression analysis

2. Multiple linear regression (10 hours)
    - Multiple linear regression, estimation and prediction
    - Regression models for quantitative and qualitative predictors
    - Model selection and validation
    - Diagnostics
    - Remedial measures
3. Logistic regression, Poisson regression and Generalized linear models (8 hours)
    - Logistic regression, inference, model selection, tests for goodness of fit, diagnostics, prediction
    - Poisson regression
    - Generalized linear models
4. Design and analysis of single-factor experiment (4 hours)
    - Experimental studies, observational studies, and causation
    - Experimental studies: basic concepts
    - Overview of standard experimental designs
    - Single-factor ANOVA model

Assessment

Homework	20—30%
Mid-Exam(s)	30—20%
Project	20%
Final Exam	30%

Syllabus by: Lulu Kang and Despina Stasi, 02/05/2018

MATH 4XX/546 Introduction to Time Series

Course Description: This course introduces the basic time series analysis and forecasting methods. Topics include stationary processes, ARMA models, spectral analysis, model and forecasting using ARMA models, nonstationary and seasonal time series models, multivariate time series, state-space models, and forecasting techniques.

Enrollment: Elective for B.S. in Statistics, B.S. in Applied Math, M.S. in Applied Math, Ph.D. in Applied Math, and all the professional master programs of the AMATH.

Textbooks: Brockwell, Peter J. and Davis, Richard A. (2002). *Introduction to Time Series and Forecasting,* 2nd edition. Springer-Verlag, New York.

References:

1. Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (1994). *Time Series Analysis: Forecasting and Control*, 3rd Edition, Prentice Hall, New Jersey.
2. Chatfield, C. (1996). *The Analysis of Time Series*, 5th edition, Chapman and Hall, New York.
3. Shumway, R.H., Stoffer, D.S. (2006). *Time Series Analysis and Its Applications (with R examples)*. Springer-Verlag, New York.
4. James D. Hamilton (1994). *Time Series Analysis*, 1st Edition, Princeton University Press,
5. Galit Shmueli and Kenneth C. Lichtendahl Jr (2016). *Practical Time Series Forecasting with R: A Hands-On Guide,* 2nd Edition, Axelrod Schnall Publishers.

Software: R

Prerequisites: MATH 475 with min. grade of C

Objective:

1. Students will learn about important time series models and their applications in various fields.
2. Students will be able to formulate real life problems using time series models.
3. Students will be able to use statistical software to estimate the models from real data, and draw conclusions and develop solutions from the estimated models.
4. Students will learn to use visual and numerical diagnostics to assess the soundness of their models.
5. Students will learn to communicate the statistical analyses of substantial data sets through explanatory text, tables and graphs.
6. Students will learn to combine and adapt different statistical models to analyze larger and more complex data.

Lecture Schedule: Two 75-min sessions per week.

Course Outlines

1. Introduction (2 hours)
    - Examples of time series
    - Stationary models and autocorrelation function
    - Estimation and elimination of trend and seasonal components
2. Stationary Process and ARMA Models (5 hours)
    - Basic properties and linear processes
    - Introduction to ARMA models, properties of sample mean and autocorrelation function
    - Forecasting stationary time series
    - ARMA(p, q) processes, ACF and PACF
    - Forecasting of ARMA processes
3. Spectral Analysis (3 hours)
    - Spectral densities
    - Time-invariant linear filters
    - The spectral density of an ARMA process
4. Modeling and Forecasting with ARMA Processes (5 hours)
    - Preliminary estimation
    - Maximum likelihood estimation
    - Diagnostics
    - Forecasting
    - Order selection
5. Nonstationary and Seasonal Time Series Models (4 hours)
    - ARIMA models
    - Identification techniques
    - Unit roots in time series
    - Forecasting ARIMA models
    - Seasonal ARIMA models
    - Regression with ARMA errors
6. Multivariate Time Series (3 hours)
    - Second-order properties of multivariate time series
    - Estimation of the mean and covariance
    - Multivariate ARMA processes
    - Best linear predictors of second-order random vectors
    - Modeling and forecasting
7. State-Space Models (3 hours)
    - State-space representations
    - The basic structure model
    - State-space representation of ARIMA models
    - The Kalman Recursions
    - Estimation for state-space models
8. Forecasting Techniques (3 hours)

- The ARAR algorithm
- The Holt-Winter algorithm
- The Holt-Winter seasonal algorithm

9. Estimation of time series models (2 hours)

Assessment

Homework	20—30%
Mid-Exam(s)	30—20%
Project	20%
Final Exam	30%

Drafted by: Lulu Kang and Tomasz R. Bielecki 02/01/2018