In the late fall of 2008 through the early spring of 2009 the Committee on Majors reviewed a proposal from the Department of Statistics to establish a new major in Statistics. This would replace the current concentration in Statistics offered within the Applied Mathematics Major. This detailed proposal included convincing rationales for the need of a stand-alone undergraduate major in Statistics and documented the growing interest in the courses offered by the Department of Statistics over the past decade. Moreover, a strong case was made that a Statistics major would provide a "diversity and plentitude of career opportunities available to the graduates of Statistics programs"*. As part of the review process, we interviewed the Chair of Statistics, Joseph Chang and Lisha Chen, a likely DUS for the proposed major. We were also provided with numerous letters from current and recently graduated students strongly endorsing the establishment of a major in Statistics. In both the proposal and in response to our queries, it seems that establishment of this major will not negatively impact the existing Applied Mathematics Major which has shifted its emphasis to areas of quantitative analysis "less explicitly statistical". Based on this review, The Committee on Majors recommends that the Yale College Faculty approve the creation of the Major in Statistics.

Curricular Design: The proposed program, appended below, will provide students training in basic concepts and methodologies of Statistics. It will address "core theory of Probability and Statistics" as well as emphasize training in methods of computational analysis, visualization of data, "methods of seeking and establishing structure and trends in data, and a mode of questioning and reasoning that quantifies uncertainty".

Resources: The establishment of the major will require very few new resources in terms of facilities, faculty, administration or establishment of new courses. The Department of Statistics has a robust and comprehensive menu of courses in place that provides the platform for the major program outlined below. They are populated with a faculty adequate in numbers and dedicated to the teaching these courses.

In conclusion, although the COM is charged with restraining the growth of new majors, in our view, the establishment of a major in Statistics seems timely given the strength of the Statistics faculty and the overarching importance of statistical analyses, including the development of new methodologies, to so many different disciplines that require rigorous, quantitative assessment of data.

Respectfully Submitted,

Mark Mooseker (Chair)
Kyle Farley (Secretary)
George Levesque (ex officio)
Joseph Gordon (ex officio)
Shannon Craigo-Snell
Hannes Leeb
Mitchell Reich, PC '09

*Note that text in quotes has been excerpted from the Statistic proposal presented to the Committee.

Proposed Major Requirements for Statistics:

The proposed curriculum for the new major is a synthesis of theory, methods, and applications. It contains a set of required and elective courses covering the fundamental concepts, theory, and methods in Statistics and Probability as well as their applied aspects.

Systematic development of Probability and statistical inference involves rigorous mathematical treatment, and working with data requires at least basic computing skills. Therefore, the prerequisites consist of mathematical preparation, and a course in computing is required in the major. The requirements are designed to achieve some balance and depth in each of the three directions of Probability, Statistics, and data analysis, by including

- two courses in the theory and applications of Probability: Probability Theory (Stat 241a) and Stochastic Processes (Stat 251b)

- two courses emphasizing the theory of statistical inference: Theory of Statistics (Stat 242b) and Linear Models (Stat 312a)

- two courses in the methods and practice of data analysis, chosen from the four options Introductory Data Analysis (Stat 230b), Data Analysis (Stat 361a), Multivariate Statistics for Social Sciences (Stat 363b), and Case Studies (Stat 625a[#1]).

Beyond the above, the BA degree requires 3 more Statistics courses. Two of these 3 courses are electives. The third course is the Senior Seminar and Project (Stat 490a), taken in the fall of the senior year. In Stat 490a, the students will work on a research project, present and share their progress with each other during the seminar meetings, and write a final report. For students who have become very interested in the topic of their senior project and wish to continue working on it further, there would be the possibility to continue the project as an independent study with the supervision of a faculty member in the spring. The BS would add additional depth by requiring one additional Statistics elective and one course in mathematical analysis.

For elective courses, there are many possibilities that can be attractive and beneficial for undergraduates. In addition to the choices indicated in the data analysis section of the major requirements, further undergraduate courses that we offer every year include a Bayesian course in Probability and Statistics (Stat 238a), Advanced Probability (Stat 300b), Information Theory (Stat 364b), and Data Mining and Machine Learning (Stat 365b). Several graduate level Statistics topics classes are offered each year, and many of these will be perfectly accessible to our majors, and so can be taken by qualified students with permission. And several courses offered by the Biostatistics Department are appropriate, such as Survey Sampling (BIS 538b), Fundamentals of Clinical Trials (BIS 540a), Applied Regression Analysis (BIS 623b), Categorical Data Analysis (BIS 625b), Longitudinal Data Analysis (BIS 628b), Applied Survival Analysis (BIS 630b), Theory of Survival Analysis and Its Application (BIS 643a), and Nonparametric Statistical Methods and Their Applications (BIS 646a).

Thus, a brief description of the requirements for the major that we would envision for the YCPS copy would look like this:

- Prerequisites: Multivariate Calculus (MATH 120a or b or equivalent) and Linear Algebra (MATH 222 or MATH 225).

- Number of courses: BA – 10 term courses beyond prerequisites (including the senior seminar and project); BS – 12 term courses beyond prerequisites (including the senior seminar and project).

- Specific courses required: Probability Theory (STAT 241a), Theory of Statistics (STAT 242b), Stochastic Processes (STAT 251b), Linear Models (STAT 312a).

- Distribution of courses: For the BA, a course in computing (ENAS 130b or CPSC 112a or b); two courses in the methods and practice of data analysis chosen from Introductory Data Analysis (STAT 230b), Data Analysis (STAT 361a), Multivariate Statistics for Social Sciences (STAT 363b), and Case Studies (STAT 625a); and 2 additional elective Statistics courses numbered above 200. For the BS: the same as BA, plus one more elective Statistics course numbered above 200 and a course in mathematical Analysis (MATH 260a or 300b or 301a).

- Senior requirement: Senior project. Seniors work on a research project and write a report, taken as a senior seminar (STAT 490a).

- Substitutions permitted: With permission of the DUS, STAT 238a may be substituted for STAT 241a, Case Studies (STAT 625a) may satisfy the computing requirement, and appropriate courses in other departments may count toward the major.

Our requirements are comparable with those by peer institutions. For a BA degree, in addition to the prerequisites in mathematics and a course in computing, we require 8 Statistics courses at intermediate level or above plus a senior project, as compared to 9 courses plus a senior seminar at Harvard, 8 courses at Columbia and 6 courses plus a year-long project at Duke.[#2]

Sample Courses of Study

Students who consider pursuing a Statistics major early in their undergraduate study are encouraged to take the prerequisites by the end of their freshmen year and start to take core statistics courses in their sophomore year. Below is a example course of study for a BA for such a student.

Freshman Year
Calculus (Math120a or b)
Linear Algebra (Math 222a or b)
Computing (ENAS 130b)

Sophomore Year
Probability (Stat 241a or Stat 238a)
Theory of Statistics (Stat 242b)
Stochastic Process (Stat 251b)

Junior Year
Data Analysis (Stat 361a)
Linear Models (Stat 312a)
Multivariate Statistics for Social Sciences (Stat 363b)

Senior Year
Senior Project (Stat490a) 2 elective courses

Such a schedule also has plenty of room in which to complete the additional requirements for the BS degree.

Some students might realize their interest in Statistics rather late in their undergraduate study. These students are still encouraged to consider the major as long as they take the calculus classes by the end of their sophomore year. They might intensively take courses in their junior year, especially in the spring semester. Below is a sample of course study for such a student.

Freshman/Sophomore Years
Calculus (Math120a or b)
Computing (CPSC 112a or b)
Introductory Data Analysis (Stat 230b)

Junior Year
Linear Algebra (Math222a)
Probability (Stat241a or Stat 238a)
Theory of Statistics (Stat242b)
Stochastic Process(Stat 251b)
Multivariate Statistics for Social Sciences (Stat 363b)

Senior Year
Linear Models (Stat312a)
Senior Project (Stat490a) 2 elective courses

[1] We intend to add an undergraduate course number to Case Studies, since we feel it is will be of interest to our undergraduate majors and suitable for them.

[2] Details on requirements of Statistics majors at Harvard, Columbia, and Duke can be found at:
http://webdocs.registrar.fas.harvard.edu/ugrad_handbook/current/chapter3/statistics.html
http://www.college.columbia.edu/bulletin/depts/stat.php?tab=dept
http://www.stat.duke.edu/programs/undergrad/major.html