

## STATISTICAL RESEARCH–TRAINING–CONSULTING AT A MATHEMATICS RESEARCH CENTER (CIMAT, MEXICO)

Fernando Avila and Miguel Nakamura  
Centro de Investigación en Matemáticas, Mexico  
nakamura@cimat.mx

*Centro de Investigación en Matemáticas (CIMAT) is a federally funded mathematics research center in Mexico, founded in 1980. It is organized into three academic departments: mathematics, computer science, and probability/statistics. With research as a backbone, its mission also includes training at the master's and doctoral levels in probability and statistics, and—much more recently—providing statistical consulting services to external clients. Provisions were taken to incorporate consulting as an essential activity at CIMAT, including a new office to serve as liaison between clients and researchers. Statement of missions, guidelines for academic faculty, and bylaws were also revised. To foster interaction between researchers, students and consulting project leaders, a virtual internal organization—the Statistics Lab—was conceived. We will review our notion of the Stat Lab, and summarize our experiences in exposing graduate students to consulting problems and in bringing together students and researchers. Lessons learned regarding an academic institution involved with professional consulting will be discussed.*

### INTRODUCTION

The subject of statistics undoubtedly rests strongly on applications. In teaching the subject, a relevant question is how to divulge statistical foundations and to simultaneously involve students in applications. By applications, we adopt a broad meaning, which includes mathematical modeling, multidisciplinary interaction, statistical thinking, computing literacy, prescription and implementation of statistical methods, as well as interpretation and communication of results. Many of these skills are not attainable through standard text-book courses, and imply some sort of complementary activities, special courses, or a general a state-of-mind.

In this paper, we will describe the evolution of teaching statistics in the graduate program at Centro de Investigación en Matemáticas (CIMAT), in particular related to applications. We will do this by describing the evolution of statistical consulting as a whole at CIMAT, and its impact on teaching and research, in the hope that some lessons can be learned and that many issues are raised for discussion. Whereas some institutions may now be asking how to incorporate consulting into the statistics program, the question we mainly address here is this: Given that consulting is a recognized activity and a well-developed structure for consulting is already established in an academic institution, how should one maintain its viability and incorporate applied aspects into the teaching of statistics?

CIMAT is one of the 21 scientific research centers not in Mexico City that belong to a Federal system, under the umbrella of the education and science and technology ministries. It was founded in 1980 in a context of decentralization, by a modest group of four researchers in mathematics, and from the beginning, was involved with training students at the undergraduate level at Universidad de Guanajuato. CIMAT has since grown to 86 academic personnel and 44 staff, and 282 students (172 undergraduate, 70 master's, 40 doctoral). The General Director is appointed by the Director of CONACYT (Consejo Nacional de Ciencia y Tecnología) and is overseen by a Governing Board with representatives from several Federal and State authorities. Its annual budget is currently 89.7% obtained from federal funds and 10.3% from funded projects and contracts.

With research as a backbone, CIMAT's mission also includes training at the masters and doctoral levels in probability and statistics, mathematics, and computer science, and—much more recently—providing statistical consulting services to external clients. This explicitly balanced mission is a distinguishing feature, as compared with universities (where emphasis is on teaching), purely research centers (where emphasis is on research), and technological centers (where emphasis is on applications). This, together with favorable historical circumstances for re-

organizing and re-defining guidelines for academic faculty and bylaws has allowed CIMAT to experiment with novel notions of consulting, and its relation to the teaching of practical statistics.

The Department of Probability and Statistics at CIMAT is an eclectic mix of theoretical probabilists and applied statisticians. Graduate programs cover subjects in probability theory, statistical theory, and applied statistics, and content design and teaching are typically handled by each of the two groups disjointly. Consulting in the department has mainly focused on statistical topics, reflecting diverse areas of expertise (sampling, time series modeling and prediction, linear models, likelihood methods, spatial statistics, design of experiments, applications to biology, statistical modeling, and others). Recently, many topics in applied probability have been explored by the group of probabilists (risk theory, branching processes in genetics, probability models in finance, and others), but no *bona fide* consulting problems in probability have yet occurred.

### CLASSIFICATION OF CONSULTING PROBLEMS

A simple device we call the “upscaling pyramid” has proved to be quite useful at CIMAT for understanding the nature of consulting problems and for explaining relationships with teaching and with research. The idea is to distinguish four levels of problem complexity and technical sophistication of a required solution. At the very bottom are “sub-professional problems,” involving very standard statistical techniques, of the type that a statistics student would be able to fully implement and prescribe. A step above are “semi-professional problems,” that require some form of deeper understanding and involvement, typically addressable by master’s in statistics graduates. A third level, “professional problems,” includes situations whose identification requires more-extensive modeling experience, and the solutions involve non-standard procedures, perhaps published in the statistics literature. The top level, “research problems,” include more-complex problems that involve applied or pure research.

This pyramidal structure (Figure 1) is used to depict two aspects. The first is that at any given time, the numbers of consulting problems coexist and decrease within each type and maintain a healthy balance. On the other hand, clients who enter the pyramid at a lower level are naturally inclined to resubmit a consulting problem at the next upper level if the consulting relationship was positive. This ascending force is a consequence of a greater degree of confidence and experience gained exercised by a satisfied client.

The pyramid can also be explained in terms of the client profile and how he/she sees the statistical consultant. At first, the lower level, the consultant is seen as a simple provider of solutions to simple problems; success at this makes the consultant a trusted source. When a consultant is perceived as a capable and trusted source then long term and more complex problems can be discussed. The highest level is when the consultant is called upon by the client as a partner to discuss strategic planning, which may include a research agenda.

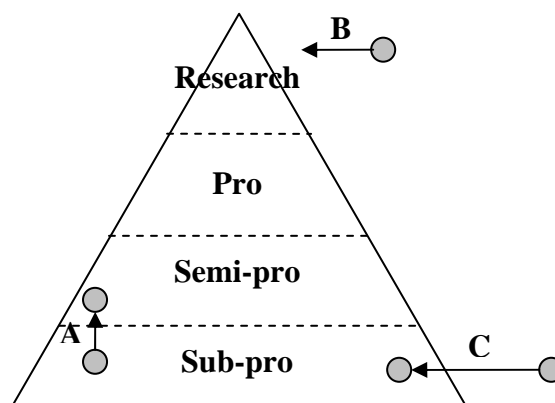


Figure 1: The “upscaling pyramid”

This simple device explains many stories of success and failure. For example, although level four is in fact the type of problem that a research center is most interested in becoming involved with, it is not at all natural for a first-time client to approach CIMAT for a consulting problem at level four (Arrow B). Dismissing new problems merely because they would be located on the first step of the pyramid (Arrow C) is not sensible. This is not only because students in training have a better chance of hands-on experience regarding problems of level one, but because over time, a satisfied client will tend to rise in the pyramid when given a chance.

If a University or a research center is recognized as being top level, then it is possible that some clients may skip some steps and require top level collaborations, either through grants or through the establishment of consortiums. Graduate students are hired, usually as research assistants, to work in these projects, but their participation is not part of their formal education although the research experience can be invaluable.

#### TEACHING AND CONSULTING AT CIMAT: EVOLUTIONARY STAGES

The evolution of the Statistics Laboratory has been intimately intertwined to the growth and evolution of the basic activities defining the CIMAT mission. For ease of presentation we can define four stages and trace the concepts and implementations in each stage.

##### *The Early Stage (1980s)*

The first group of researchers in statistics came to CIMAT and they established a master's program in statistics. Since they viewed statistics from the "methods" point of view the master's program was heavily oriented to traditional applicable courses. Since there was no experience in statistical consulting, students were only exposed to textbook applications. Consulting services were personal and informal and included students only if a thesis could be developed. There was no Statistics Laboratory. Our upscaling pyramid was in fact, completely empty at this stage.

##### *The Middle Ages (early 1990s)*

There were changes in the researchers' roster and research groups started developing and the first externally funded research projects were carried out. There was a shift in the master's program toward a more theoretical and rigorous orientation. However, some multidisciplinary interactions began which fostered some thesis with an emphasis toward applications and this was the only applied training that some students had. Since these thesis topics were driven by research reasons, some of these thesis were in fact level three or four, but the clients were still personal contacts and purely academic in nature.

On the other hand, paid consulting services were provided by a business unit which was created solely for that purpose and did not have any ties to the research groups or to teaching. The projects that were developed were straightforward and short term statistical and SPC applications, that is, on level one and with no forces to drive problems up the steps in the pyramid. The business unit did not function as a Statistics Laboratory, mainly because there was no academic component.

##### *The Renaissance (late 1990s)*

Researchers were added and grouped according to areas of interest. There were several externally funded research projects. The diversified interests led to the creation of two separate master's programs, the traditional one with a theoretical flavor, and a new program with an emphasis in quality control and industrial statistics.

The first consulting contracts were obtained and CIMAT recognized the importance of creating a formal structure to foster and support consulting services. The first Statistics Laboratory was created and support staff with a master's degree in statistics was hired.

CIMAT formalized the structure for consulting services through the creation of the Technological Services office.

A course of statistical consulting was created and was used to expose students to case studies, thesis topics and a discussion of consulting as practiced by the project leaders.

The first successful consulting contracts were in the lower levels of the pyramid, but the clients returned and have kept a higher level relationship with CIMAT, eventually producing some papers and doctoral theses.

*The Current Era (2000s)*

The number and complexity of consulting projects has steadily increased in the last five years and there has been a renewed interest and success in multidisciplinary interactions. There was a need to create the position of statistical consulting manager and to reformulate the Statistics Laboratory concept in terms of a consulting services model, which was adopted by CIMAT when it recognized consulting as an activity in equal terms to basic research and teaching.

The master’s programs were unified as a unique program having a common backbone and a series of optional terminal courses. Statistical consulting is one of the required courses. Students can and have developed thesis as part of consulting projects.

The pyramid concept has been validated because now CIMAT has strategic clients that began in the lower levels of the pyramid and are now regarding CIMAT as a trusted adviser and have kept long term contracts. The goal of having 10% of the Center’s budget coming from consulting services was reached last year and will be reached again this year.

INTERACTIONS AT CIMAT

Interactions between research (R) and teaching (T) have always been natural in the academic environment, and have been so at CIMAT since its creation. Thesis topics and involvement of students as research assistants are self-driven. However, the notion of a Statistics Lab (SL) and the relation of consulting (C) with R & T are fairly new. For example, the model depicted in Figure 2 is used in many universities as a way of strengthening statistical education.

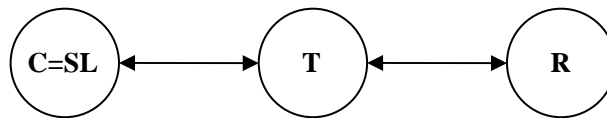


Figure 2: A simple model for teaching-consulting interactions

Consulting here is viewed as a mere complement to teaching. The SL is conceived as a place for users of statistics to come discuss problems, and possibly even have a solution implemented or sketched by students. There are no fees charged, but a sub-professional solution would be norm. Interactions with users of statistics (“clients”) are seen as ways of enriching the learning experience for statistics students. Perhaps some academic interdisciplinary collaborative work would evolve from this, but no fully professional contracts. This was the first conception and implementation of a Stat Lab at CIMAT.

The evolutionary stages of interactions described in the previous section have thus resulted in an organization at CIMAT that reflects the three main activities in its mission: R, T, and C (Figure 3). SL is a virtual entity that stands at the center of a triangle, serving as a channel of information and a repository. Each stated node is the connection with a class of external individuals who conduct business with CIMAT, and each has a well-identified head (the Department Chair for R, the Director of Graduate Studies for T, and the Director of Technological Services, for C). The SL connects all vertices within CIMAT, enabling relevant matters between any pair of nodes. For example, whereas in early stages of development C was viewed completely detached from T and R, it is now viewed as relevant both to R (complex consulting problem may require research component; researcher may be willing to participate in consulting) and T (students can participate in simple consulting problems; consulting problem may employ student labor during the summer). Notice that in fact, there is a buffer between a consulting client and the SL, so that the SL is no longer viewed as being there for the sole purpose of consulting.

Direct impact of this organization on teaching is twofold. On one hand, consulting problems have a channel for reaching students and students and instructors have a channel for

reaching problems, including an external user of statistics, a context, and data. On the other hand, students can reach research problems and researchers can readily access the pool of students for purely academic matters. The latter impact is usually vigorous and natural in academic institutions, but the former one regarding consulting problems is not generally made explicit.

Students are encouraged to participate in projects. They can do it in several ways. They are regularly exposed to ongoing projects as part of their training, either in consulting classes or in seminars. The consulting course itself, required for all students in the master's program, includes non-technical aspects of consulting such as communication skills, the writing of reports, and negotiation techniques, as well as practical and foundational aspects of statistical modeling. As part of the consulting course they are required to participate in at least one project, writing a report about the statistical problems and a proposed solution. There are also field trips to industries or research centers to get a direct view of the problems exposed by the clients.

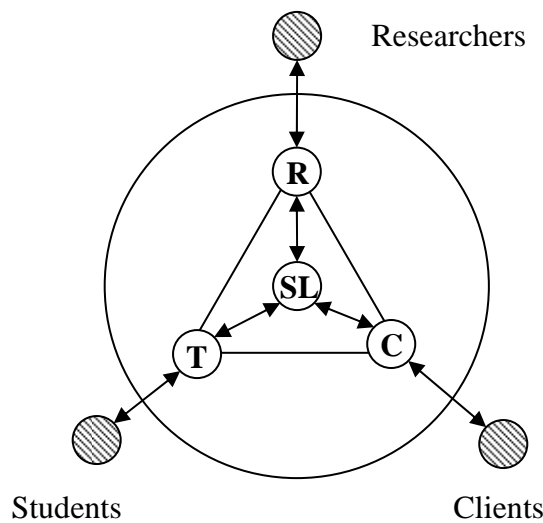


Figure 3: Three-way interactions at CIMAT

Students can also participate in a project as part of the team. In this case it can lead to a thesis project, or it can be as technical support for specific tasks. In both cases, they receive monetary compensation, either as a scholarship or as wages charged to the project. When a student participates in a project it is expected that his/her work follows the rules and requirements of the project.

It is too early to judge if this type of statistical training has a positive impact on future jobs and/or performance of our students. The first generation of students that shared the consulting experience (course, field trips, project exposure) is one semester old. What is definitely true is that a great number of prospective students are attracted to our program when they learn about this infrastructure.

## DISCUSSION

The key challenge we have faced is to keep the Stat Lab as an equilibrium point in the triangle R–T–C. There are several threats and opposing factors to this. We must be aware of them and maintain a continuous struggle against them. We list some of the current issues at CIMAT.

Success increases demand, and consulting services have been steadily increasing, a situation that is coupled with internal and external pressures to increase revenues. However, the roster of researchers and especially of those who participate in consulting has been kept constant, partly due to federal constraints. The temptations are a) accept only those projects that are easy (lower steps in the pyramid) and/or generate a large income, b) saturate up to a burn-out the teams doing consultation services and c) improvise with students and researchers that are not ready for quality consulting. The sad truth, applicable to most academic settings, is that some researchers will never do consulting either because they don't want to, or because they can't.

Success has put the need for a tight organization. CIMAT has issued a set of guidelines, rules and regulations, flow diagrams, role description, etc. As with any other bureaucratized structure, there is the apparent dilemma of either going by the book, and become slow and inefficient, or to skip steps and “cut corners” in order to meet deadlines. There have been instances of “staff grabbing,” of “project grabbing” and of “unjust compensation,” partly due to the resistance of the researchers to “swim with the current.” The fair assignment of students to projects and these advisors requires the joint ruling of the graduate committee and the SL, as a prevention of “student grabbing.”

Success has also glamorized consulting activities and students are strongly attracted to this, declaring that they only want “applied” courses. When CIMAT opened a second master’s program with this orientation, some, and among them the weaker students, thought that this was a weaker program and in fact turned it into a low grade version of the original “method’s” master’s program. This situation also caused tensions and strong discussions among the teachers until a decision to re-unify into a single program was reached.

When a student chooses to do a thesis around a project there is the question of balancing between satisfying the client and satisfying the thesis committee. This is not an easy situation and it can lead to a disappointed party with the student having spent months of hard labor and caught in the middle. Not all projects have problems with enough statistical merit for a master/doctoral thesis, although the client may think so; a joint committee may be a way of combining both points of view but only if the client has a strong academic background to understand these requirements.

Our discussion has been restricted to the master’s program in statistics, but CIMAT also hosts a doctoral program and it hasn’t been clear what role can doctoral students have in our consulting structure. This is due to several factors. First of all, the main emphasis on doctoral studies is research with high academic quality, applied work is suitable only if it is at the highest level of the pyramid and these types of projects may be scarce or untimely. Doctoral students coming from other institutions are not trained or encouraged to do applied work and/or come focused to do basic research.

Some students that have the master’s degree from CIMAT have been hired as technical staff for specific periods of time, and have gone on to doctoral studies abroad, choosing a program that consolidates the vision acquired during their job as consultants. There is a natural tendency to stay at CIMAT, either at a hired job or as a doctoral student, but this may be a case of diminishing returns and therefore is not generally encouraged.

## CONCLUSIONS

We have presented a model for the consultation services provided by a research center. In this model a Statistical Laboratory is an integrating entity joining research, teaching and consultation.

In this paper we emphasized on the impact of the model and the structure put in place at CIMAT on the teaching and training of students. We described both the advantages and the key challenges that we have experienced. Since this a new conception, we still don’t have a definite evaluation, but it has been an exciting and promising start.

## REFERENCED DOCUMENTS

The ubiquitous relationships between consulting, teaching and research at CIMAT can be perceived throughout the following documents.

Current bylaws for academic faculty at CIMAT (in Spanish) may be found at [http://www.cimat.mx/pag\\_externa.php?m=2&ext=tra\\_index\\_epa.htm](http://www.cimat.mx/pag_externa.php?m=2&ext=tra_index_epa.htm). To note here is that consulting is recognized firsthand as a substantive activity, and that faculty and research assistants may focus on this.

Regulations designed to govern income and expenses related to consulting activities are described at [http://www.cimat.mx/pag\\_externa.php?m=2&ext=tra\\_index\\_rec\\_prop.htm](http://www.cimat.mx/pag_externa.php?m=2&ext=tra_index_rec_prop.htm).

General guidelines for the workings of the laboratories described above can be read at [http://www.cimat.mx/pag\\_externa.php?m=2&ext=tra\\_lab\\_vinc.htm](http://www.cimat.mx/pag_externa.php?m=2&ext=tra_lab_vinc.htm).