Statistical education for doing Statistics professionally: some challenges and the road ahead

Pedro Luis do Nascimento Silva

Escola Nacional de Ciências Estatísticas (ENCE/IBGE)
Outline of Talk

• Who will be doing Statistics Professionally?
• How do we ‘teach’ each group?
• Statistical evolution and challenges
• Opportunities
• Some ideas and experiences
Who Will Do Statistics Professionally: Three Major Groups (Paths)

- **‘Professional’ Statisticians**: those who get a first degree in Statistics, who may or may not get further graduate education in Statistics.
- **‘Graduate’ Statisticians**: other professionals who will get formal statistical education only during post-graduate studies.
- **‘Self-taught’ Statisticians**: other professionals who will get statistical education through their work, but not through formal statistical education.
Professional Statisticians

• People who
  – Like “Math”
  – Choose career at an early age
  – Have limited knowledge of what the career involves, because it not as well known as Engineering or Medicine
Educating of Professional Statisticians

- **Consequences**
  - Low demand for places offered (around 5 candidates to 1 place here in Brazil)
  - Large drop out during undergraduate Statistics courses
  - In Brazil, only about 1/3 of those which start the course get the degree
## Brazil – Some Statistics On Undergraduate Programs

<table>
<thead>
<tr>
<th>Indicator</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Courses</td>
<td>24</td>
<td>25</td>
<td>26</td>
</tr>
<tr>
<td>Places offered</td>
<td>1.331</td>
<td>1.423</td>
<td>1.464</td>
</tr>
<tr>
<td>Places occupied</td>
<td>1.154</td>
<td>1.269</td>
<td>1.145</td>
</tr>
<tr>
<td>Total enrolment</td>
<td>3.837</td>
<td>3.941</td>
<td>4.507</td>
</tr>
<tr>
<td>Graduates</td>
<td>353</td>
<td>266</td>
<td>390</td>
</tr>
<tr>
<td>Graduates/Places occupied</td>
<td>30.6%</td>
<td>21.0%</td>
<td>34.1%</td>
</tr>
</tbody>
</table>

Source: Census of Higher Education - INEP
Undergraduate Programs in Statistics - Brazil

Graduates / Places occupied (%)

Source: Census of Higher Education - INEP
Professional Statisticians: Basic Curriculum 1970s

- Calculus and Linear Algebra
- Descriptive Statistics (no EDA)
- Scientific Computing (FORTRAN)
- Probability Calculus
- Statistical Inference
- Design of Experiments
- Linear Regression
- Multivariate Analysis
- Time Series Analysis
Professional Statisticians: How We Were Taught -1970s

- Mathematics and Mathematical Statistics dominated the curriculum
- Emphasis on how to compute or to obtain statistical output, often “by hand”
- A students’ consumption dream in the 1970s
- Little time spent on the analysis and interpretation of the outcomes of the statistical work
Changes Of The 1980s

- Mainframe computers
- ‘General’ statistics packages (SPSS, SAS, BMDP)
- In comes EDA
- Mid 80s see the arrival of
  - GLIM, CART, SPAD
  - Plus many other specialized statistical software
- Accessibility limited to government agencies, main universities and large companies
Changes of the 1990s

- Microcomputers
- Spreadsheets and the end of ‘manual’ statistical calculations
- PC statistical software (S & others)
- Statistics get into small businesses, even into the households
- Graphical displays and interactive data analysis
- Nonparametrics, smoothing
- Bootstrapping
- Bayesian Statistics and Simulation
Most Recent Changes

• Internet ‘everywhere’
• Google
• Google Scholar
• R
• Who knows what comes next
‘Graduate’ Statisticians

• People who
  – Get first impressions of Statistics during undergraduate studies and/or at work
  – Get interested in Statistics via applications
  – Convert to Statistics career after first degree in Engineering, Economics, Mathematics, Computer Science, even Medicine

• Learn Statistics for application in their areas of work / interest
Graduate Statistical Education – Type 1 – ‘Traditional’ Orientation

- Probability Theory
- Stochastic Processes
- Mathematical Statistics
- Statistical Modeling (linear and nonlinear)
- Multivariate Analysis
- Time Series
- Then more specialized courses + thesis
Graduate Statistical Education – Type 2 – Applications Orientation

- Biometrics
- Demography & Social Statistics
- Econometrics
- Operations research
- Process quality and control
Type 2 ‘Graduate’ Statisticians: How Do They Learn Statistics?

• Data and problem driven presentation of methods and techniques ➔ application-oriented curriculum

• Limited time spent on theory

• Specialized statistical ‘language’ for each field

• ?Do we see them as Statisticians?

• ?Do they see themselves as Statisticians?
Impact Of Evolving Technology

• Graduate statisticians
  – Dominate subject matter in application area
  – Less time spent in statistical calculations

• Professional statisticians
  – Increased opportunities for statistical work and collaboration
  – Need to stay competitive ➔ permanent effort to learn and keep track of statistical developments
Some Key Questions

• What is important to know for doing Statistics professionally?
• What are the challenges of today’s and tomorrow’s statistical practice
Statistical Challenges (1)

- Data are messy, abundant and everywhere
- Data must be organised, processed, analysed and interpreted to be useful
- Ever increasing set of statistical tools (theory, methods, software) to obtain and analyse data
- Increasing demand for use of Statistics in almost every branch of human knowledge
Statistical Challenges (2)

• Increasing complexity of problems and applications
• Increasing risk of misuse of Statistics, by users having access to sophisticated statistical tools, yet lacking of sufficient statistical competence
Francisco Cribari Neto (2002)

• “Statistics biggest challenge is to promote the notion that Statistical Theory, Statistical Practice and Computational Statistics cannot be separated”

• Facing this challenge requires upgrading statistical education, as offered to both Professional and Graduate Statisticians
Core Competence & Skills Required To Do Statistics Professionally (1)

• Solid knowledge of statistical principles and methods, as well as of their properties and limitations

• Capacity to recognize, formulate and resolve real problems involving data collection and analysis, applying such principles and methods in adequate manner
Core Competence & Skills Required To Do Statistics Professionally (2)

- Capacity to collect, organise, process, analyse and present data using computing tools, fast and accurately
- Be critical and able to detect situations where existing methods do not work and need to be adapted or new ones developed
Core Competence & Skills Required To Do Statistics Professionally (3)

- Effective communication both orally and in writing
- Working effectively in multidisciplinary teams
- Working effectively with existing technology
- Develop an attitude to learn continuously
Are We Meeting the Challenge?

• Are we educating people and professionals with all these “core” competences and attitudes?
• If not, why?
• What are the weaknesses of our standard Statistical Education programs
• And what do we do to “keep in shape” after we complete our formal education?
Some Barriers

• Rigid curricula
• Emphasis still reflects mathematical side of Statistics, without sufficient effort to develop the other required competences
• Lots of “deductive teaching”, with proofs and demonstrations – mechanical, uninspiring
• Pedagogical approach is often old-fashioned
• Teaching of Mathematics subjects dissociated with the teaching of Statistics subjects
John Nelder, 2005

“There are some people who want Statistics to be a part of Mathematics, which it can never be.”
More Barriers

- Limited use of modern learning tools and resources
- Low prevalence of using real and relevant data and problems to present theory and develop practice
- Little emphasis on aspects of how to obtain data of good quality, and on methods to deal with data deficiencies
  - E.g.: missing data; detection and handling of outliers; measurement errors; study design;
Some Opportunities

• Ever more powerful resources available
• Evolution of statistical theory & methods
• Open source, extensible, portable and free statistical computing tools available
• Abundance of data and examples of relevant applications available to motivate and provide “learning by doing” opportunities
• Facilities available for distance education and learning, and cooperation
• Receptivity to the use of Statistics
Estela Bee Dagum (2005)

• “Students must learn statistics by ‘doing’, ‘writing’ and ‘talking’ ”.
• “Teachers should be able to motivate their students by making classes more attractive”. 
Some Ideas to Move Forward

• Reorganise curricula for greater flexibility and balance regarding development of key competences
• Encourage introduction of new pedagogical approaches
• Increase use of real data and problems as learning tools
• Increase use of R, the internet, Google
Some Ideas to Move Forward

• Incorporate fieldwork type projects as part of the learning curriculum
• Lead by example: Statistics teachers MUST do Statistics
• Mentoring and apprenticeship programs also to be made part of the curriculum
• Commitment with continued education throughout ones’ career, after graduation
How We Can Improve

• More research in statistical learning and education
  – ICOTS, IASE, etc.
• More cooperation
• Motivating and educating TEACHERS for using new approaches and technologies
• Give STUDENTS greater control over the learning process
• ➔ Perhaps E-learning could help here
How We Can Improve

• Quality pre-university statistical education introduced to provide basic statistical literacy even before university

• Greater involvement in projects and multidisciplinary teams

• Creating options for Statistics as complementary to other university courses
Some Things We Tried At ENCE for Education Of Professional Statisticians

- Two semester credit units devoted to project development by teams of students
- Projects are required for graduation
- Students choose project topic and have one or two teachers as project supervisors
- Several projects relate to problems brought from students’ workplace
Some Things We Tried At ENCE for Education Of Graduate Statisticians

- MSc. program providing statistical education for those interested in Demography and Social Research
- Survey skills development course
- Capacity building at ENCE to enter the Distance Learning era
  - Team devoted to run training courses for IBGE staff by means of distance learning
  - One member of staff on leave doing PhD in subject to return soon
Electronic Distance Learning (EDL)

- A powerful new approach for teaching and learning, Statistics and many other things
- It takes advantage of new information and communication technologies
- It makes room for a more personalized learning experience, by allowing students or apprentices to make progress at their own pace, and to study anywhere, anytime
Electronic Distance Learning (EDL)

- Within this new approach, students must be more pro-active in the learning process.
- EDL replaces ‘traditional classrooms’ by ‘virtual learning communities’, using communications devices which enable remote synchronous and asynchronous interactions between teachers and students, or just between students.
- Content dissemination may be enhanced by including interactive ‘materials’ that improve presentation and facilitate learning.
Electronic Distance Learning (EDL)

- Global costs of learning may be lower, after the initial investment on the required technology
- Opportunities are created for better integration of content preparation by different teachers
- Provides means for continued education both at the workplace or at home that would otherwise be unavailable with current approaches
EDL Weaknesses

• It may lead to fewer interactions between students and teachers
• Group learning is more difficult
• It may reduce student motivation sources and opportunities
• Those without technology and communications skills and resources may be inhibited from participating, both teachers and students
EDL Weaknesses

• Prejudice is still high about this type of teaching and learning approach
• Self-discipline is required from students
• If learning is to happen at the workplace, adequate conditions and time for study must be provided at the workplace too
• Assessment and certification are more challenging
EDL – Critical Factors

• Although technology is often considered the essential component that must work well, to succeed EDL also requires first-rate content preparation coupled with adequate pedagogical treatment.

• Hence EDL projects are very demanding in terms of the skills required to their development and execution, requiring important changes to teacher’s approaches to preparing and providing content, as well as about their role in the learning process.
EDL – Critical Factors

• Perhaps the difficulties in adapting to take advantage of this new approach explain why we still have so few examples of comprehensive projects in Statistical Education carried out by means of EDL

• But the size of the challenge should not discourage ourselves

• The larger are the challenge and the risks of the enterprise, the greater are the expected rewards
EDL – Critical Factors

• We should consider ways to work together to develop this new approach and bring it into the mainstream of statistical education
  – At university (both undergraduate and postgraduate) and professional levels (including continuing education at the workplace)
  – In pre-university development of statistical literacy

• I encourage you to consider when and how you might get involved with this fascinating new world of EDL in statistics
Official Statistics Agencies

• Can help statistical educators by providing access to data and other relevant support material for teaching

• Portugal: ALEA project – aims at providing both teachers and students of secondary education with teaching materials for the study of Statistics

• http://alea-estp.ine.pt/
Official Statistics Agencies

• Canada
• E-STAT:
  – Interactive learning tool designed to support the education community
  – Offers access to warehouse of statistics about Canada and its people
• http://www.statcan.ca/english/Estat/intro.htm
Official Statistics Agencies

• Brazil
• Learning support for children and teenagers
  – IBGE Teen
• http://www.ibge.gov.br/ibgeteen/frameset_open.html
  – IBGE 7 a 12
• http://www.ibge.gov.br/7a12/default.php
Above All

- Develop an attitude to learn continuously
- And cooperate with others to achieve that
Thank you for the attention!
Enjoy ICOTS7 and Salvador!
HP 67 Calculator

Source: http://www.hpmuseum.org/3qs/673q.jpg
HP97 Calculator

Source: http://www.hpmuseum.org/97.jpg
HP 67 Calculator

Source: http://www.hpmuseum.org/67.jpg
IBM /370-148 (1976)

Source: http://www-03.ibm.com/ibm/history/exhibits/mainframe/mainframe_2423PH3148.html
IBM-PC (1981)

Source: http://www-03.ibm.com/ibm/history/exhibits/pc/pc_1.html
Evidence for Need to Learn Continuously

• How many operating systems have you had to learn ever since you started to use computers for statistical work?
• How many text processors have you learned to use to process your texts?
• How many statistical software packages have you had to learn throughout your career?
• How many new statistical techniques have you had to learn to improve upon your practice?