# TEACHING STATISTICS TO PHYSICIANS: A FIVE-YEARS EXPERIENCE

Ângela Tavares Paes
Universidade Federal de São Paulo, Instituto Israelita de Ensino e Pesquisa Albert Einstein, Brasil
atpaes@unifesp.br

*Physicians are often exposed to papers and academic theses that involve complex statistical analyses. Due to the development of computational facilities, they are able to produce many p-values, but rarely are worried about the adequacy of the methods employed. Even though, there is a big demand for statistical courses among medical researchers and postgraduate students who are interested in perform their own analyses and better understanding scientific papers. In this context, traditional courses in which students are required to compute standard deviations or search critical values in theoretical distributions tables do not make sense. The biggest challenge on teaching statistics to postgraduate students in medicine is to focus on interpretation from a practical point of view without forget or put the theory in second place. This paper describes some difficulties and alternative methods to teach statistics that have been used in introductory courses at a public university and a private hospital.*

## INTRODUCTION

The use of statistical methods in medical research has grown exponentially over the past 20 years and become strongly consolidated in biomedical investigation. Currently, all scientific papers involving data analysis employ statistical inference and we rarely find papers with descriptive analysis only. In medical meetings or journals, physicians are bombarded with statistics. Besides the difficulty in understanding lectures and published papers, they also face many difficulties when trying to publish their own studies in scientific medical journals, which have become more rigorous in their judgments regarding to the quality and adequacy of statistical analysis employed in the submitted papers. These difficulties involve not only those with poor statistical knowledge, but also those who have a good background despite of the lack of knowledge of more recent methods. As a consequence, physicians are urged to do statistical courses that may help them to survive in the academic and professional world. More than an urgent need, it is a survival issue for medical researchers.

Under this scenario, the statistical courses have emerged as an important way, almost mandatory, for clinicians who want to keep actualized and active. In Brazil, although basic-statistics courses have proliferated, many teachers are not sensitive enough to understand the difference between postgraduate physicians and undergraduate students, therefore, many of them insist in traditional courses in which students are required, for example, to compute standard deviations or search critical values in theoretical distributions tables. This does not work for medical researchers who are frequently exposed to statistics and need to known how to interpret analyses published in medical papers.

In this paper I discuss difficulties associated with teaching statistics for clinicians as well as some purposes of changes. As background, I consider my experience in introductory courses for postgraduate students at Federal University of Sao Paulo and Albert Einstein Hospital.

## OUR BASIC-STATISTICS COURSES AND SOME STATISTICAL ISSUES

The first version of the course offered at Universidade Federal de São Paulo (UNIFESP) was elaborated by Professor Clovis de Araújo Peres in 1996. His purpose was to reformulate the conventional way of teach, with emphasis on intuition rather than mathematical formulations. It is a dangerous choice, because there is a high risk of deleting important theoretical issues extremely necessary for understanding the main concepts. When you try to give a more practical way the failures can happen. In UNIFESP we have several examples of courses taught by practitioners who commit serious conceptual errors.

Peres's idea was successful and his formula has been adopted in the last fourteen years by some teachers (and by me). Some changes have been made depending on the teacher, but the essence of focus on interpretation remains unchangeable.

In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php [© 2010 ISI/IASE]

The course does not require prior knowledge and concepts are introduced through practical exemplification. It is a one-semester course with 12 classes once a week. The main difference when compared to conventional courses offered to undergraduate students is related to the way we use to introduce statistical tests. In a general setting, we call of "measure of distance from the null hypothesis" to refer to statistic's tests (the quantity computed from data to be compared to a critical value). The theoretical distribution that provides critical values is referred as "reference system". If the "measure of distance from the null hypothesis" computed is higher than values obtained from a "reference system", it means the null hypothesis is not plausible and must to be rejected. The p-value works as a translation of this comparison. It is worth to mention that these students are used to see and compute p-values, despite of their poor knowledge about statistical methods. The computation of p-values is explicitly presented only as an example for a one-sample z test. Instead, we discuss the flexibility of p-values interpretation and we recommend the students must not rely solely on p-values because clinical relevance is more important than statistical significance.

Based on the general rule about "measurements of distances" and p-values, the most common statistical tests are introduced as solutions to evaluate relationships among variables. For this, it is essential that students know perfectly how to classify variables into categorical or numeric. In that way, Chi-square test is used to study the association between two categorical variables, t tests and ANOVA can be a way to exploit the relationship between a categorical variable and a numeric dependent variable. When the interest is to analyze relations between two numeric variables, Pearson's correlation coefficient and simple regression models can be an approach. Figure 1 shows the course contents. We can note that some topics such as design of experiments and probability theory were omitted.

Regarding to design of experiments, we realize that it would be impossible to cover all the alternatives planning designs in a short-time course. Besides, most of our students have already defined the design which is, in general, a one-center randomized or an observational study. Also, sampling planning strategies are not discussed, since almost all of studies they work use convenience (not probabilistic) samples. Therefore, we assume that the design has been defined, data was collected and we replace the design of experiments topic by tips of how to organize databases.

Concepts of probability theory are introduced intuitively and not as an isolated topic. As the students have a minimum background about probability from undergraduate courses, they do not have difficulties to understand probabilities associated to hypothesis testing and statistical modelling. Also, with exception of Normal distribution all the other theoretical distributions are seen as "reference systems" to be used for testing hypotheses.

We do not adopt any text book, and the slides presented in class are available by e-mail. Obviously, we emphasize to students that this material can not be seen as a book and its purpose is to help them in monitoring and better understanding the classes. The criterion for approval is based on performance on two written tests. The student is approved if the weighted average of the two tests is at least 5 points (on a scale of 0 to 10). In the first years, the approval was based on a practical work related to the student's thesis. The idea was to link their work with concepts presented in class. Although this is interesting and useful for them, many students just "copy-paste" texts that they have already written before and were not worried in understanding what we have discussed during the course. Written tests were a suggestion given by them and we agree it would be more effective.

The course is offered 2 times per year and currently there are only two classes with 60 seats each. Unfortunately we are unable to attend the demand. Last semester, there were about 300 interested students, therefore 140 students failed to enroll.

Figure 1. Contents of basic-statistic course offered at UNIFESP

- Introduction (definition of Statistics, the interaction between statistics and the scientific method), classification of variables and organization of databases.
- Descriptive analysis: how to describe categorical variables - frequency tables, bar charts and sectors.
- Descriptive analysis: how to describe numeric variables - measures of position and dispersion, histograms and box-plots.
- Introduction to statistical inference - standard error, sampling error, Normal distribution, confidence intervals and hypotheses tests.
- Associations between two categorical variables - Chi-square, Mc Nemar, Kappa.
- Relationship between a categorical variable and a dependent numeric variable – tests for means comparisons - Student t test, paired t test and ANOVA.
- Relationship between two numeric variables - Pearson's correlation and simple linear regression.
- Influence of a set of (categorical and numeric) variables on a numeric normally distributed variable - Multiple linear regression.
- Influence of a set of (categorical and numeric) variables on a binary variable - Logistic Regression.
- Special Topics: survival analysis, evaluation of diagnostic tests (sensitivity, specificity, predictive values, ROC curve) and non-parametric tests.

At Hospital Albert Einstein the course is offered to graduate physicians and follows the same ideology than UNIFESP courses. One important difference is that the students are not necessarily involved with theses yet and the pursuit of knowledge is not as urgent as is in UNIFESP. On the other hand, the physicians are almost daily asked to read and discuss scientific papers; therefore, they have more contact with complex methods and new methodologies. We present some theoretical and practical classes and, in addition, we have "journal clubs" where we discuss recent papers published on medical journals.

## STUDENTS' OPINION

The course at UNIFESP has several limitations and has to be considerably improved. Even though, according to students' opinions it is very useful for their practice. In 2006, we applied a survey satisfaction with 139 students who have completed the course. The students were asked to attribute a score from 0 to 10 related to the following aspects: utility, duration, facility to understand the concepts and teacher's performance. In Table 1 we can note that, in general, the course was well evaluated with averages higher than 8 in all aspects. The smallest average for duration can be explained by the need of the students of learning more. This is a short course where the topics are discussed superficially. In their opinion, time is not enough to exploit all the topics. Many of them suggest a longer time or a second course with advanced statistics.

Figure 2 shows some spontaneous opinions given by the respondents in the survey and by students who concluded the course in the subsequent semesters. Fortunately, these comments are frequent and prove that, in spite of the limitations, we are going in the right way.

Table 1. Course evaluation–results from a survey with 139 students–UNIFESP, 2006

| Aspects | Mean | Standard deviation | Minimum | Maximum |
|---|---|---|---|---|
| Utilility of the course for you | 9,46 | 0,97 | 7 | 10 |
| Duration (hours) | 8,43 | 1,74 | 5 | 10 |
| Facility to understand the concepts | 8,68 | 1,77 | 4 | 10 |
| Teacher's performance | 9,86 | 0,36 | 9 | 10 |

Figure 2. Comments given by students about the basis-statistics course in UNIFESP

*"The course was extremely helpful to me, even to repeat it if you have opportunity".*
*"I love these classes because they deal with methods very common in the reading of scientific articles and resources that I needed to use in my thesis. After the course, I could better understand the issues and participate actively in the statistical analysis of my thesis, including reproduce some calculations I did not understand so clearly before."*
*"For the first time, I understand the theory of logistic regression."*
*"Now I can read the medical articles without skipping statistical analysis section."*
*"The course was excellent; can I repeat it?"*
*"I lost my fear of Statistics."*
*"I want to congratulations the discipline of Biostatistics for the opportunity to offer graduate students the chance of a better understanding of Statistics in health."*

The most negative point observed in the survey was the lack of practical classes with use of computers. The students have a real need to perform analyses and more than theoretical concepts, they want to learn how to deal with statistical software. As a result of several suggestions, we have decided to include some practical classes. This was not a simple task and will be discussed in the next section.

## USE OF STATISTICAL SOFTWARE

The development of statistical software with user-friendly interfaces has allowed statistics users to perform their own analysis. Therefore, many clinicians have become independent from statisticians. From my experience as a consultant, this change was notorious. Fifteen years ago, most of clinicians were used to require statisticians to do everything, since simple descriptive analyses until more complex models. Nowadays, they are able to perform the statistical analysis and require a statistician only to do a "check list" or answer questions made by a medical journal's referee about a submitted paper.

I consider that the accessibility to statistical software can be "a powerful weapon given to children". If on the one hand these software help to disseminate the statistics and make it more important among medical researchers, on the other hand they might trivialize statistical methods and make them easier than they really are. Apart from this particular opinion, I agree that statistical courses with only theoretical classes are incomplete and do not attend the physicians needs.

The lack of practical classes was the main complain of our students when we asked them to evaluate our course. However, we have several limitations to incorporate classes with use of software. First, commercial statistical software are very expensive and would be so hard to buy a license for 60 users. Another problem is the physical space. As our laboratory is small, we would have to share the students into two or three groups and give practical classes in different days or periods. The schedule is very limited in a one-semester course with only 12 classes.

An alternative to expensive commercial programs would be to use free statistical software such as R developed by the R Development Core Team (*http://www.r-project.org*). However, some years ago the software R was exclusively a command-driven system, that is, the user had to reproduce analysis by typing commands. The requirement the user remembers the names and arguments of commands increase the chances of syntax errors, which make the use of R relatively difficult by clinicians who are not entirely familiarized with computers and statistical programs. These difficulties were significantly minimized in 2005, when the R commander (Fox, 2005), known as Rcmdr, was developed. The Rcmdr provides a basic-statistics graphical user interface to R in which simple and familiar menu/dialog-boxes are used. Last semester, we introduce a

practical class with use of R in which the students reproduced analysis in their personal notebooks. For some of them, it was difficult to install the packages, but most students have approved the experience. At the moment we are providing a more detailed guide with examples from the theoretical classes where the commands are showed step-by-step.

## ESSENTIAL DIFFICULTIES

Besides the difficulties and limitations mentioned previously, we consider some essential points to be discussed.

- Absence of a text-book. For many teachers it would be impossible or very hard to teach without a text book. Although I agree with this idea, I am not able to choose a single book to use in my course. A book I consider excellent was written by Douglas Altman. Even though, the course is based on a list of books and texts written by myself. Also, I often use Altman's short articles published at British Medical Journal.
- Use of statistical software. Statistics-courses are frequently confounded with software-courses. Unfortunately, many physicians have the wrong thought that it is enough know how to operate a statistical program to perform statistical analyzes. Currently, we have to deal with this trivialization of statistics. Most software cannot assist in the understanding of statistical principles. Nevertheless, I believe that courses with only theoretical classes are not successful among clinicians.
- Clinician's view about statistics. One important difficulty associated to teaching statistics to medical practitioners and researchers is that the vast majority expect to find quick recipes to get the desired results, not understanding the complexity of some statistical methods. It is quite common to see requests or attempts to standardize approaches which sometimes cannot be standardized because depend on some subjective aspects. We have to show them that data interpretation must be flexible and take into account the researcher's knowledge.

## CONCLUDING REMARKS

The demand for courses in statistics among professionals in the health field is very large. The biggest challenge on teaching statistics to medical researches is to focus on interpretation from a practical point of view without forget or put the theory in second place. According to my experience and my students' opinions I believe it is possible, but we have to deal with some difficulties such as the urgent clinician's need to apply statistical methods in their research and constant exposure to data analysis and user-friendly statistical software.

Despite some notable successes in biomedical studies and improvements of researcher's knowledge, the misuse of statistics in the medical literature remains common. Well designed statistics courses are extremely important to minimize inadequate analyses present in published papers and improve the quality of scientific research.

## ACKNOWLEDGEMENTS

I would like to thank Clovis de Araújo Peres and Lisbeth K. Cordani who have introduced me to the activity of teaching statistics and for their great contribution to education in Brazil.

## REFERENCES

Fox, J. (2005). A Basic-Statistics Graphical User Interface to R. *Journal of Statistical Software*, *14*(9), 1-42.

Altman, D. G. (1991). *Practical Statistics for Medical Research*. London: Chapman & Hall.