A system of student-specific practical sessions of statistics on the Internet

Miñarro, Antonio
Sánchez, Josep Anton
Arcas, Antoni
Vegas, Esteban
Reverter, Ferran
Calvo, Miquel

University of Barcelona. Department of Statistics
Diagonal 645
Barcelona (08028) Spain
E-mail: aminarro@ub.edu

1. Introduction and objectives

The authors are members of *Statmedia group*, a team of professors of the Department of Statistics at the University of Barcelona (UB). Our department has the responsibility of teaching statistics in a variety of degrees of our campus: Biology, Geology, Environmental Sciences, Chemical engineering, Geography and Statistics. So we have a wide experience in teaching statistics both for applied and more theoretical courses. In many experimental degrees, Statistics is an instrumental learning often offered in the first or second year. The aim of our group is developing new tools to help teaching statistics. The system has been implemented mainly in experimental courses but we have also work on it with more specialized students of the degree in Statistics. Our project is based on some important assumptions:

- First, we consider that learning should be based on practical cases to motivate and encourage students to participate.
- Second, the system should be applicable no matter how many students are registered. Some of the tasks the system must be capable of include are: continuous assessment of knowledge based on a large number of tests, follow-up of student's attendance, possibility of making different parts of the course away from the classrooms but under the supervision of the teacher, etc.
- Third, our system should be independent of any commercial software that involves any cost for the students.
- Finally, our main goal is improving academic performance of our students and to enhance motivation towards statistics.

Those requirements have been implemented in a system of design, publication and management by Internet of different sets of practical sessions of statistics with specific data sets for every student of the course (Statmedia 2PI). Customization of data is the most notable achievement of the project since we consider it an essential requirement to allow students not to be present in the classroom. It also expands the possibility of discussion since different students may have different results on different data sets under similar assumptions. Even different questions may be generated depending on the data sets.

Previous tests with identical data for all the students showed a negative attitude from many students that simply copied results from other companions. That practice invalidates the results as a continuous assessment procedure.

One of the main challenges of this approach is developing a system of automatic and student-specific correction of the answers allowing for a wide variety of questions: closed item tests, numerical open questions and combined questions with several answer fields.

In our present project we have integrated a previous work in the same field. The edition of a multimedia textbook of statistics with more than 750 html pages and more than 100 java applets edited by the University of Barcelona on a cd-rom format in 2003. The main part of the contents of the cd-rom is now accessible from the pages of the practical sessions.

2. Technical matters

The framework of the project is based on three elements: JSP pages, databases on a web server and applets and servlets of java language.

On the server side, a small set of servlets manage the entire application, tracking the main parameters of each connected session. This includes login validation, database transactions and navigation across the documents. The navigation is performed by a servlet which calls a JSP, following the user requests and the configuration of the practical case. Any called JSP sends to the client a HTML document, build dynamically when the server inserts the individual data of the student inside a pre-formatted page. If necessary, the server also configures particular questions and pre-load data for the applets embedded in the document. JSP pages are easily formatted by our design team, who has only to include some *Statmedia* special tags (similar in syntax to the JSP specification) when editing the HTML document. A second JSP set is designed only for internal use, helping the teacher to interact with the database stored information, in particular, querying for useful summaries of the student's answers and activities.

A Java language statistical tool to help in computations has been developed by our team. The format is that of a statistical calculator that is called by a button control. It appears as an independent window offering all the statistical computations required by the practical sessions. Since the whole system has been developed in Java it is possible for the calculator to interact directly with the JSP pages. In this way different data sets can be automatically loaded in the calculator by means of a load button. Student's time and mistakes can be substantially reduced. Figure 1 shows the aspect of a part from a practical session with several kinds of questions and a window corresponding to the statistical calculator. You can see several item tests and some open fields.

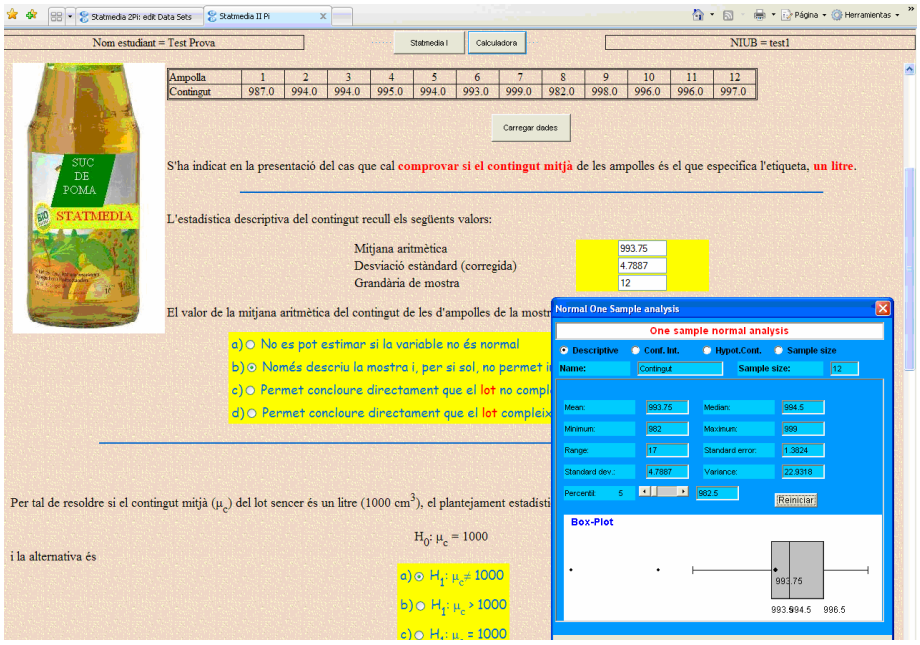
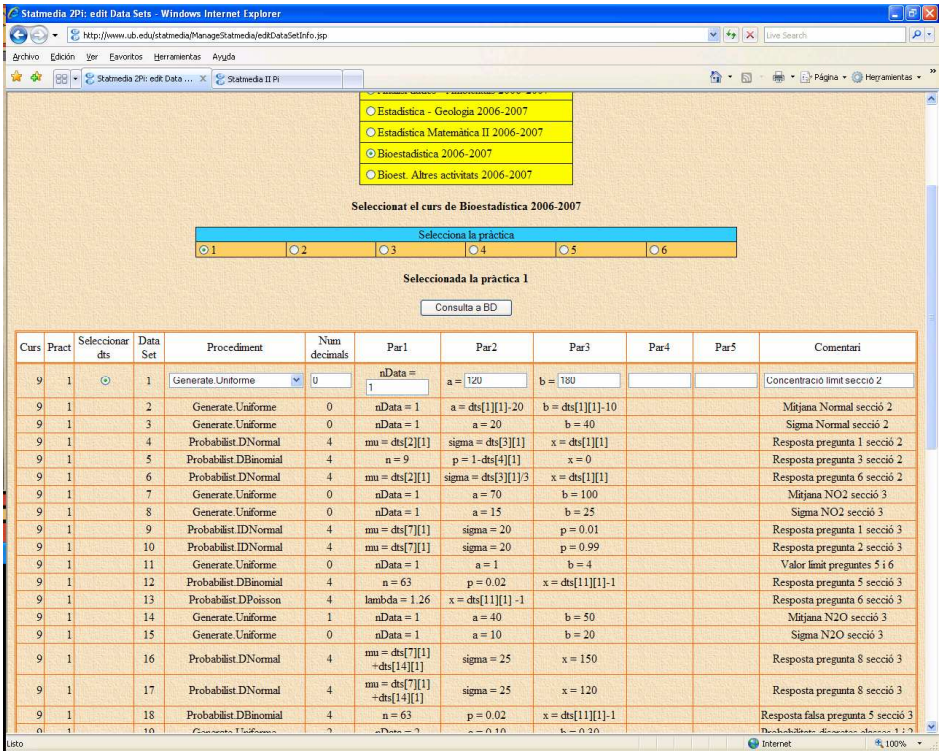


Fig. 1 Aspect of a page corresponding to a practical session with the statistic calculator and some questions of different kind

Student-specific data generation is obtained through the definition of what we call "*dts*" (data sets). A *dts* is an array of values generated by specific java method. Such methods include: random numbers generation from different distributions; computation of critical values and distribution functions; descriptive statistics for one sample; one and two sample analysis with test statistics, p-values and other computations; categorical data analysis; correlation and regression for two samples and finally transformations and operations between data sets. Random numbers generation methods allow student-specific data sets. It is important to note that the system just stores the definition of each data set, not the data. The system works with a seed depending on a private identification of the student. This procedure assures the consistency of data at any time.



Curs	Pract	Seleccionar dts	Data Set	Procediment	Num decimals	Par1	Par2	Par3	Par4	Par5	Comentari
9	1		1	Generate.Uniforme	0	nData = 1	a = 120	b = 180			Concentració límit secció 2
9	1		2	Generate.Uniforme	0	nData = 1	a = dts[1][1]-20	b = dts[1][1]-10			Mitjana Normal secció 2
9	1		3	Generate.Uniforme	0	nData = 1	a = 20	b = 40			Sigma Normal secció 2
9	1		4	Probabilist.DNormal	4	mu = dts[2][1]	sigma = dts[3][1]	x = dts[1][1]			Resposta pregunta 1 secció 2
9	1		5	Probabilist.DBinomial	4	n = 9	p = 1-dts[4][1]	x = 0			Resposta pregunta 3 secció 2
9	1		6	Probabilist.DNormal	4	mu = dts[2][1]	sigma = dts[3][1]/3	x = dts[1][1]			Resposta pregunta 6 secció 2
9	1		7	Generate.Uniforme	0	nData = 1	a = 70	b = 100			Mitjana NO2 secció 3
9	1		8	Generate.Uniforme	0	nData = 1	a = 15	b = 25			Sigma NO2 secció 3
9	1		9	Probabilist.IDNormal	4	mu = dts[7][1]	sigma = 20	p = 0.01			Resposta pregunta 1 secció 3
9	1		10	Probabilist.IDNormal	4	mu = dts[7][1]	sigma = 20	p = 0.99			Resposta pregunta 2 secció 3
9	1		11	Generate.Uniforme	0	nData = 1	a = 1	b = 4			Valor límit preguntes 5 i 6
9	1		12	Probabilist.DBinomial	4	n = 63	p = 0.02	x = dts[11][1]-1			Resposta pregunta 5 secció 3
9	1		13	Probabilist.DPoisson	4	lambda = 1.26	x = dts[11][1] -1				Resposta pregunta 6 secció 3
9	1		14	Generate.Uniforme	1	nData = 1	a = 40	b = 50			Mitjana N2O secció 3
9	1		15	Generate.Uniforme	0	nData = 1	a = 10	b = 20			Sigma N2O secció 3
9	1		16	Probabilist.DNormal	4	mu = dts[7][1] +dts[14][1]	sigma = 25	x = 150			Resposta pregunta 8 secció 3
9	1		17	Probabilist.DNormal	4	mu = dts[7][1] +dts[14][1]	sigma = 25	x = 120			Resposta pregunta 8 secció 3
9	1		18	Probabilist.DBinomial	4	n = 63	p = 0.02	x = dts[11][1]-1			Resposta falsa pregunta 5 secció 3

Fig. 2 Definition of "datasets" for practical session number 1 of Biostatistics

Methods that compute statistical results allow JSP pages to offer different dynamic texts based on different results. For instance you can ask different questions if you know whether two samples of one student are correlated or they are not. You can also take into account the result of different tests in order to offer the student dissimilar answers. It is important to note that the designer of the practice can control, via the *dts*'s definition the fraction of students with each and every different result. Next we can see several lines of code where the system shows a value for a specific data set (dts[2][1]) and different alternatives for a question depending on the result obtained by data set dts[16][4].

```
<td ><font size="4"><b><%= dtsFormat(dts[2][1],0)%> </b></font> </td>
…..
<%      if (dts[16][4] < 0.05){ %>
            <% ind.getPermString( "We reject the null hypothesis",
                                  "We accept the null hypothesis",3)%><% }
        else { %>
            <% ind.getPermString( "We accept the null hypothesis",
                                  "We reject the null hypothesis",3) %><% }%>
```

A further outstanding feature is that automatic assessment of student's answers is carried out by the same java methods used by the statistical calculator. In this way we are sure of the consistency of the results obtained by the student and the ones from the automatic evaluation system.

3. Some results

In recent years over 700 students per year in four different first and second cycle studies have used Statmedia 2Pi. We have several tools to analyze the degree of satisfaction and the academic achievement of students. Some of the results of courses finished by now are next summarized.

At the end of the year, a survey is passed out so that students can express their opinions about several aspects of the project. We may relate some answers with the main assumptions we have considered important in section 1:

- The convenience that learning be based on practical cases has been scored with 4.66 over 5.
- The agreement with student-specific data and the improved results discussion has been scored with 3.47 over 5.
- The possibility of making different parts of the course away from the classrooms has been scored with 4.58 over 5.
- The continuous assessment procedure has been scored with 4.64 over 5.
- The global satisfaction index is 7.4 over 10.

As we have mentioned above JSP technology has provided us with several tools to analyze student's performance. Through those tools we have verified some facts:

- About 67% of registered students have actually done the practical sessions of statistics. It represents a better one of that obtained in other similar courses, for instance in mathematics.
- 35% of total editions were found to be done on computers no belonging to the UB.

If we now consider the academic achievement of students, the average mark of practical sessions has turned out to be 8.07 over 10. In the context of continuous assessment we have implemented, practical sessions contributes 20% to the final mark of the course.

With respect to the teachers, the system is considered a very useful, dynamic and flexible tool for teaching statistics. It has been also a very helpful tool for convergence to the European Space for Higher Education

REFERENCES

Becker, B.J. *A Look at the Literature (and Other Resources) on Teaching Statistics*. Journal Educational and Behavioral Statistics, 21, (1996).

Calvo, M et al; *Statmedia: Un curso multimedia de Estadística*. Publicacions de la Universitat de Barcelona (2003).

Johnson, H.D.; Dasgupta N., *Traditional versus Non-traditional Teaching: Perspectives of Students in Introductory Statistics Classes,* Journal Statistics Education Vol.13, Number 2 (2005)

Timothy S. Vaughan T.S., *Teaching Statistical Concepts With Student-Specific Datasets,* Journal Statistics Education Vol.11, Number 1 (2003),