Enriqueciendo la experiencia del aprendizaje de Estadística con animaciones interactivas D3.js: Aplicaciones para seguros de las cadenas de Markov

Enriching the statistics learning experience with D3.js interactive animations: Insurance applications of Markov chains

Antonio Fernández Morales
Universidad de Málaga
afdez@uma.es

RESUMEN.
El objetivo de este artículo consiste en la exploración de las posibilidades que las animaciones interactivas pueden ofrecer para enriquecer la experiencia de aprendizaje de Estadística en el ámbito de la educación actuarial. Un tipo particular de animación, basado en la librería Javascript D3.js, ha sido seleccionado para ello por sus potentes componentes de visualización, así como por su tratamiento natural de las transiciones. Esta segunda característica es de vital importancia para la visualización gráfica de las cadenas de Markov. Algunas aplicaciones de las cadenas de Markov en el campo de los seguros se simulan usando este entorno y se han aplicado en el curso de procesos estocásticos del Máster en Ciencias Actariales y Financieras de la Universidad de Málaga. Finalmente, se analizan los resultados de la experiencia, incluyendo los de una encuesta realizada al final del curso, que revelan las principales fortalezas del enfoque adoptado, tal como han sido percibidos por los estudiantes.

PALABRAS CLAVE.
Animación interactiva, simulación, seguros, formación actuarial.

ABSTRACT.
The aim of this paper is to explore the possibilities that online interactive animations offer to enhance the statistics learning experience in the field of actuarial education. A particular type of animation, based on the D3.js JavaScript library, has been chosen due to its powerful visualisation components and natural treatment of transitions. The latter is especially adequate for the graphical visualisation of Markov chains. Some insurance applications of discrete time Markov chains are simulated using this visual framework and used in a course of Stochastic Processes of the MSc in Actuarial Science at the University of Málaga. Finally, the results of the experience along with the outcomes of a survey conducted at the end of the course are analysed, revealing the main strengths of this approach perceived by the students.

KEY WORDS.
Interactive animation, simulation, insurance, actuarial education
1. Introduction.
It is widely recognised that learning and teaching Statistics in Higher Education is specially challenging in the Social Sciences area. Factors like students’ scarce motivation, which generally provokes a limited participative attitude, and an insufficient contextualization of didactic resources, are frequently behind the main obstacles to overcome. Even though there is an increasing awareness of the importance of research in Statistics education, it is still considered an emerging discipline (Tishkovskaya & Lancaster, 2012).

A group of lecturers in the MSc in Actuarial Science at the University of Málaga has been working in this field since 2003. Our main objective is to get a more effective learning process in several academic fields, including Statistics education. The approach followed is to take advantage of new technologies, combined with an active learning environment according to a cognitivist strategy.

Following this quite long experience using new technological resources to enhance the Statistics education in the MSc in Actuarial Science, a new resource, based on interactive animations, has been specifically designed to enrich the learning process of a specific topic that is not particularly well received by our students: Stochastic Processes. In order to get a really effective implementation, the interactive animations simulate discrete time Markov chains contextualized in the field of Insurance.

This paper is structured as follows. The methodology section describes the main approach followed in the design of the interactive animations, including the pedagogical aspects, as well as the technological ones. In the next section, some preliminary results of a survey are analysed to know our students’ perceptions about the usefulness and effectiveness of the digital resources proposed. Finally, the last section includes the main conclusions of this study.

2. Material and methods.

2.1 Technological issues.
The new, and constantly evolving, Information and Communication Technologies (ICT) offer a wide variety of resources to use in Higher Education. It is currently acknowledged that these resources should be more than simple collections of static resources delivered through a digital channel (Monahan et al., 2007, Dinov, Christou & Sánchez, 2008). More recent proposals are focused on interactive, dynamic resources suitable for blended, online and even mobile learning.

There are many examples in Statistics education combining new pedagogical approaches with a significant technological component that aim to reinforce the learning process while increasing students’ motivation (Lundsford et al., 2006, Symanzik & Vukasinovik, 2006, Dinov, Christou & Sánchez, 2008, Liu, 2010; González et al., 2010, Wang et al., 2011, Lin et al., 2015). A well-established line of research in this area consists of integrating digital resources into the general strategy specifically designed to ease the comprehension of the most difficult or abstract concepts (Chance et al., 2007, Schneiter, 2008, Byrne et al., 2010, Sosa et al., 2011).
Although there is a wide range of this type of resources, since the nineties most of them used to be based on two predominant interactive technologies for web pages, Flash and Java applets, being the second one the most used. Many Java applets have been developed for enriching the learning process of different key concepts in Statistics, like hypothesis testing (Schneiter, 2008), the central limit theorem (Dinov, Christou & Sánchez, 2008), or confidence intervals (West & Ogden, 1998).

Following this approach, some Java applets for the Actuarial Statistics courses in the MSc in Actuarial Science at the University of Málaga have been designed and used with quite good results (autor, 2008, 2010, 2011).

Recently, a new paradigm is being adopted regarding the design and use of interactive graphical resources on the Internet. The conjunction of new capabilities of browsers with the powerful Html5, CSS, and SVG standards is encouraging the development of a different approach based on open source Javascript libraries. Unlike Flash or Java, these are non-proprietary developments, and are very well suited for mobile devices access. A recent example by the author (autor, 2014) has adopted these standards to build an interactive seasonal concentration simulator for the Tourism degree.

One of the most interesting and popular Javascripts currently available for this purpose is D3.js. It is conceived for manipulating documents based on data and also provides powerful visualisation components. Its data-driven approach to DOM manipulation along with its focus on transformation allows very remarkable animated transitions and functions to control visualisations.

The interactive animations designed for the Stochastic Processes course in the MSC of Actuarial Science at the University of Málaga are fully based on D3.js. These animations use the Markov chains development by SETOSA² (Powell & Lehe, 2015) with some additions.

2.2 Pedagogical approach.

A blended learning approach is followed, in general, in all the courses of the MSc in Actuarial Science at the University of Málaga. This hybrid strategy is based on the combination of face-to-face activities with online resources designed to enhance the learning process while potentiating the autonomous component (Mayorga-Toledano, 2010, autor, 2010), including e-learning and m-learning resources (autor 2003; autor, 2010).

A special emphasis is placed on the contextualization of all the digital resources. The rationale of stressing the importance of a highly contextualised environment comes from the necessity of translating the actuarial profession competences into the syllabus of the MSc (autor, 2015, autor, 2011). Our students are required to identify and acquire these professional competences within a didactic environment that potentiates authenticity, collaboration and creativity (autor, 2013). Authenticity is also a key factor in the assessment model we are working on since 2009 (autor, 2009).
In particular, the course in Stochastic Processes follows the approach described in the previous paragraphs and introduces several on-line activities based on an interactive graphical simulator. The interactive simulator allows students to experiment with various types of insurance applications of Markov chains. This type of resource has two appealing features to enhance Statistics learning: (i) the dynamic visualization greatly facilitates the understanding of some complex concepts and (ii) the interactive nature of the simulator allows experimenting. Many authors consider the latter as one of the most effective tools for improving the efficiency of Statistics learning (Chance et al., 2007, Littlejohn et al. 2008, and Wang et al., 2011). Moreover, this kind of simulators can be included into one of the learning objects classification of Churchill (2007).

The animated simulations are fully integrated into the general scheme of the course’s blended strategy. They are proposed to be used either as a complement in some activities taking place in the face-to-face sessions as well as in activities to be done autonomously by our students. This way, the learning process we are pursuing is not a mere sum of independent components, but a coherent progression of activities and methods (Garrison & Kanuka, 2004, Naismith & Corlet, 2006).

In addition, the possibilities and sizes of the screens of mobile devices have been taken into consideration in the design of the animations, in order to adapt the contents to a realist usage.

2.3. The interactive animations.

Several versions of the simulator have been developed to visualise the dynamic process of discrete Markov chains in insurance applications: a simulation of a basic version of the gambler’s ruin problem, a bonus-malus scheme with five classes with increasing premiums for motor insurance (autor, 2015), and a simplified model for fire insurance with three classes. They all follow the same general design (Fig.1), embedded into a web page according to the colour and font scheme of the course website.
The general context of the insurance application is briefly described on the upper part of the web page, leaving a generous space behind for the animation and controls. Although the descriptions of the applications have been deliberately kept very short, they include all the statistical details to fully understand each problem and to allow an adequate interpretation of the results. Figure 2 shows an example of one of the animations, the one corresponding to the fire insurance simplified model.

The main area is divided into four parts. The interactive sliders are placed on the upper left part. There are two sliders. The first slider controls the speed of the animation. At very low speed it is easier to observe the detailed movement of the Markov chain when needed. In contrast, sometimes it is convenient to speed up the animation to notice which states are the most frequently transited.

The second slider controls the value of the Poisson parameter in the animations that require a discrete probability model to generate the transition probabilities in the transition matrix. Changes in this slider generate real time reactions in the rest of the elements in the main area: the transition matrix elements are updated, as well as the probability distribution displayed in the extra statistical information area. In addition, changes in the Poisson parameter also modify the interactive animation in two ways: (i) changing the width of the
The visual interactive animation is shown in the left lower part of the screen. It consists of a series of elements representing the chain states, which are connected by arrows. As mentioned above, the width of the arrows is proportional to its corresponding transition probability. The movement of an animated disc represents the randomly simulated trajectory of the chain. It moves between states along the arrows, according to the real-time random simulated trajectory.

All the transition probabilities are displayed in the transition matrix on the right lower part. Every cell of the matrix is clickable. Once a cell is clicked the corresponding arrow in the interactive animation is emphasized over the rest. This is a very interesting feature that allows students to identify the elements of the transition matrix with the arrows in the state diagram. This feature of the simulator eases the understanding that both representations of the Markov chain, one more visual and the other one more technical, determine the same process dynamics.

Finally, some extra statistical information is also displayed on the upper right part, mainly the probability distribution of claims in the fire and automobile insurance applications. This statistical information also helps students to understand the construction of the transition matrix from the probability distribution of the variable which determines the process, in this case, the number of claims in a year.
3. Results.

Survey results.
In order to get a first assessment of the utility of the interactive animations presented in this paper, a survey was conducted at the end of the first course in which they were used, the course of Stochastic Processes of the MSc in Actuarial Science at the University of Málaga. Our students were asked about their perception of the potential benefits for their learning process of several characteristics of the interactive animations, with a special emphasis on those related to interactivity. The answers were recorded in a 1-5 Likert scale. Figure 3 shows the average scores obtained in the sample (bold line).
In general, the results were quite good. The overall mean scores are all above 4. The best-valued elements were the dynamics animations and the parameter control (with mean scores 4.77 and 4.82, respectively). A secondary group of remarkable elements were multi-device and on-line access, both with mean scores slightly over 4.5.
The rest of elements assessed were valued with mean scores over 4, being the lowest that corresponding to use of colours (4.32).
An additional analysis of the survey was motivated by the presence of different attitudes towards digital resources within our students. Although the answers to a question about the comfort using this kind of resources were in the range of 3-5 in the 1-5 scale, it was a clear indication of some extent of heterogeneity. Thus, Fig. 3 also shows the mean scores by familiarity with digital resources, dividing the sample into a majoritarian group of students who answered 5 to this question, 74% of the respondents, and another group with answers lower than 5 (3 or 4), that accounts for the 26% of the students in the survey.
Mean scores of the second group (red dashed line) are all lower than their corresponding ones of the first group (black dashed line), revealing that those students perceived a lower potential benefit from digital resources in general, and from this animation in particular.
However, the distribution across the different elements is very similar in the two groups. The group less prone to use digital resources also gave the highest scores to (i) the dynamics animation and parameter control and (ii) on-line and multi-device access.
The final questions in the survey were related to the intention of use, recommendation to other students and potential demand for more resources of this type in their courses. Regarding intention of use, the mean score obtained was 4.42, indicating a very good opinion of our students at this respect. A higher mean score was recorded for the recommendation question, 4.58, which reinforces the evidence of a quite good students’ perception of the usefulness of the resources assessed. Accordingly, the demand of more resources of this type can be considered very high, at least considering that the sample average in the survey was 4.63.
Interactive visualisations, combined with real time simulations, are becoming a key element in the new approaches to Statistics learning in Higher Education. On the one hand, these tools allow students to experiment with models and data in real time, facilitating the comprehension of abstract models, while easing the perception of the dynamic features of the processes, which sometimes constitute the main barrier to an adequate understanding. On the other hand, embedding these resources into web pages opens a wide variety of uses, from blended learning schemes, to MOOCs or full mobile learning courses.

The results of the survey in this investigation confirms that this type of interactive resources facilitates the understanding of Statistical concepts and models. This result can also be found in several recent studies related to Statistics education. The work of Novak (2014) concerning simulation-based learning with graduate students at a US university reveals an improvement in performance in Statistics learning when using these resources, particularly in topics like variability and the empirical rule. Dinov, Sánchez & Christou (2008) reached a...
similar conclusion in their investigation at the University of California in the US with several Statistics courses. Using interactive web-based tools for learning concepts like the central limit theorem, statistical distributions and the analysis of variance, a significantly better performance of their students was obtained.

There is also evidence of a positive effect of interactive animations on a better understanding in the field of Actuarial education in autor (2011). In that study, the surveyed students gave the highest rating in the questionnaire to the item related to easing the understanding of the most difficult concepts in the course of survival models, both for the web-based and the iPhone-based resources.

Moreover, the very good results obtained in the the survey of this research, regarding MSc students, are in consonance with the conclusions of Sosa et al. (2011), who found in their review of several worldwide investigations about computer assisted Statistical learning a positive effect on learning, greater for graduate than for undergraduate students.

Especially relevant for the development of interactive resources in this field are the possibilities offered by technologies related with open source Javascript libraries and modern browsers, particularly suited to multi-device and mobile access, which could be used to enhance the digital resources available with new capabilities.

The D3.js library is a remarkable case in this field. It provides very powerful animation and visual capabilities and takes advantage of the new Html5, CSS and SVG standards. It is also non-dependent on popular proprietary technologies like Java or Flash. In addition, there are many developments based on D3.js available, like the ones used in this work.

Particularly useful is the proposed application to visualise the dynamics of Markov chains in discrete time. Combining the animation with the possibility of controlling the parameters of the model results in an especially useful resource, as our students revealed in the survey. Both items, parameter control and dynamic animation, were the best valued ones, being an indicator of the significant pedagogical value of this application. Our students clearly stated that the possibility of experimenting in real time by changing the intensity parameter of the model was the most interesting feature for them.

Similar results regarding the advantages of interactivity and using animations have been stated by Dinov et al. (2013). In their experimental use of a web-app for the multivariate normal, conditional and marginal probabilities learning obtained an excellent response by their students. The possibility of changing the parameters, obtaining a real time reaction of the calculations and graphics, along with the rotation enabled visualizations, were the main reasons adduced by students in their assessment of the experience.

The interactive resources developed for this research are intended for use in an intermediate level of understanding and applying discrete time Markov chains, in the sense of Wang et al. (2011). These authors found the greater achievement in understanding in the intermediate levels of the four levels of interactive animations they defined (from traditional static web-based instruction to animation with interactive inputs and practice feedback) for Introductory Statistics learning in a US University, especially for the most difficult concepts related to hypothesis testing.
In our case, the visual animations followed a deliberate simple design with the strictly necessary informative elements. In addition, changes in any model element generate a real time reaction in the visualization and the information displayed on the screen. Thus, there is only one model visualization on the screen at a time. This approach is in line with the proposals Lin et al. (2015). These authors found a better performance when showing an individual interactive model than the cases of multiple simultaneous ones. Their work was based on interactive correlation simulations in courses at the National Taiwan Normal University.

Moreover, the technological component, based on a framework that allows an easy access and compatibility, has been also very well received by our students. The on-line and multi-device availability items were given high scores in the survey, thus being within the most relevant features of the interactive digital resources developed in this project.

Although the simulator was presented by the lecturer in an in-class demonstration, students made an intensive use by their selves, for personal study and also for preparing groups projects. This is an interesting feature that adds value this kind of resources. Their versatility relies on being available on a multiplicity of devices and, at the same time, being very easy to use and apply. Thus, students with different learning styles should use them in several ways: as a complement in face-to-face classroom sessions, as a resource in their personal study time, or as a resource to include in their collaborative projects presentations, to enrich them with interactive and dynamic elements.

Therefore, the design of our interactive animation resources allows both the presentation and the practice modes mentioned by Novak (2014) in his revision of applications of computer simulation methods for Statistics learning.

Using interactive simulations also resulted in a more enjoyable learning experience in our investigation. Several studies also confirm this hypothesis in the field of Statistics learning. Novak (2010) found out in his study that students preferred simulation based learning to other modes, like traditional lectures, assignments or teamwork, in terms of their enjoyment. A similar result was stated by Dinov, Sánchez & Christou (2013), based on their students’ perception of ‘more interesting’ classes than any other previous ones in the same field. This perception was also found in autor (2014), who mentioned that one of the best valued items in his survey was the ‘less boring’ feature of the interactive animations used for the Gini index learning in the field of tourism seasonality.

To conclude, it is worth noting that the applications to the study of discrete time Markov chains in insurance described in this paper have been, in general, very well valued by the students of the MSc in Actuarial Science at the University of Málaga. And thus, the results of this study confirm that an adequate use of interactive animations and simulations can enrich the Statistics learning process, as perceived by students.
5. Notes.
1 This work has been supported by Project PIE015-159: ‘Estrategias para la adquisición y desarrollo contextualizados de competencias profesionales’, Universidad de Málaga: Proyectos de Innovación Educativa, Convocatoria 2015-2017.
2 The author whishes to acknowledge Lewis Lehe and Victor Powell, for granting access to their SETOSA’s JavaScript library.

6. References.


