

Implementing Auto-generated Power Point Slides for Demonstrating Statistical Concepts

Jen-Ting Wang and Sen Zhang
Department of Mathematics, Computer Science, and Statistics
State University of New York College at Oneonta

Introduction

Proper use of visual aids can facilitate students' learning. This is especially true in teaching statistics, which often involves large amount of data that are tedious to calculate and abstract concepts that are difficult or even impossible to picture in our heads. Java applets or statistical software packages such as Fathom are available for demonstration purpose, but they have certain limitations. Microsoft (MS) PowerPoint (PPT) is one of the most popular and ubiquitous presentation tools used in classroom; however, statistics instructors seldom have access to high-quality statistics algorithm PPT slides due to the extreme difficulty of creating them by hand. To remove this barrier, we have adopted a slides generation approach to automate the tedious production process of PPT slides. The resulting generator software can be reused to automatically create new teaching slides for different input data collected by instructors. A demonstration of the Central Limit Theorem (CLT) using this technique is presented in this paper. The approach is practical, standardized, effective and innovative.

Motivation

Algorithm visualization research has always had a strong appeal for the undergraduate education community. Many educators and researchers have addressed the topic, resulting in many flavors of proprietary visualization tools or Java Applets. Most Java-based demonstrations are rendered online, rather than offline, so they cannot be saved; if the person who created an applet stops providing the applet for any reason, end users will cease to have access to demonstrations. Finally, Java is a generic programming language, rather than a dedicated presentation language; to use Java to create a presentation is therefore less efficient than using a presentation environment such as PPT for the same purpose. Professional statistics packages like SPSS or Minitab can also be used in classroom, but they were not designed for presentation purpose and will only show the final results. For teaching purposes, sometimes we not only need to show the statistical results but also the calculation procedures, which often are repetitive and tedious. The novelty of our proposed approach lies in the fact that the slides are not created directly by hand; instead, generators were implemented to automatically produce the slides. The slides will show calculations or results from each of the repetitive procedures that instructors wish to show. This approach frees instructors from tedious hand preparation of slides that demand accuracy and repetition, but the generator, a new piece of software, must be coded by programmers for each algorithm, which may be programmer labor intensive. However, the effort can easily be paid off since the generator software is highly reusable as additional instructors adopt the slides generated by the project's generators. From a pedagogical perspective, this proposed approach promotes preparedness, standardization, consistency, and the reuse of teaching materials and experiences.

Use of Technologies

The following Microsoft tools, languages, environments and technologies will be intensively used in the proposed project: Microsoft Visual Studio, C#, Office auto APIs (more specifically, PPT and Excel auto APIs), and Windows operating systems (XP, Vista, or 7).

An Example

As one of the most important concepts in statistics, the Central Limit Theorem (CLT) describes the behavior of the sampling distributions of mean and provides one of the foundations for inferential statistics. However, very often students find it too abstract to understand, which in turn challenges instructors' teaching. A typical and efficient way to explain the CLT is demonstrating it through bootstrapping. Bootstrapping requires the instructor to take multiple samples from a set of data, make numerical computations, then finally draw graphs of the distributions and observe the differences. These simple calculations sometimes need to iterate hundreds or thousands of times under various configurations. The whole procedure is extremely tedious to do manually. Software such as Fathom can show final results; but it requires acquisition and the knowledge of using it. There are some online applet examples available, but sometimes it is not possible to use the applets to show results from user collected data. Additionally, due to limited class time, it is difficult to use statistical software packages to efficiently explain the concept. Therefore, we feel the CLT would be a good topic to adopt this slides-auto-generation approach to create computer visual demonstration, which not only displays the theorem in a step-by-step manner but also is much more convenient for presentation purpose.

In order to use the technique, an instructor only needs to go through the following simple steps:

1. Collect your own data and specify the minimum and maximum sample sizes, and the number of simulations, then save as a .xml format file. These editing operations can be easily done using any plain text editors such as Notepad.
2. Run the batch file containing the source program codes under DOS command. Specify the name of the input file prepared in the above step and the Excel file and the PPT file in which you wish to keep track of the calculation, results and graphs.
3. As the generator runs, simulation results and graphs will be produced in a sheet for each sample size under the designated Excel file. Each graph produced in Excel will be shown on the automatically generated PPT slides from the designated file. For example, if you plan to show the simulation results/graph for each of the sample size from 1 through 50, there will be 50 PPT slides automatically produced and showing the results/graphs that you wish to show.
4. Instructors can post process the generated slides and spreadsheet documents. For example, they can add extra comments or highlight some results. This is an optional step, and instructors may find it very useful because of the editing ability of PPT, which is not supported by all other approaches.
5. The generated slides and Excel spreadsheets (with or without post-editing) can be repeatedly used in teaching.
6. If an instructor needs to use a different data set for any reason, then she/he can start from step 1.

Generalization

This novel technique can also be used in other statistical tests that require iterations of computation, such as chi-square test, Fisher's Exact Test, or permutation tests, etc.

Reference

S. Zhang and J. Ryder. (2012), “Using and Developing Automatically-generated Microsoft PowerPoint Slides to Facilitate Teaching of Data Structures and Algorithms “, *Journal of Computing Sciences in Colleges*, 27(3).