Full Mathematical Power In Calculated Questions Through Spreadsheets

Hiram Bollaert

Department of Business Studies, Artesis University College Antwerp, Belgium,
hiram.bollaert@artesis.be

Abstract

Moodle understands about 42 mathematical functions ranging from "+" to "tanh" and at one point we managed to get a 1171 characters long formula with nine variables working. Iimagine building your questions in a spreadsheet which has a much richer variety of functions available. With that power, vast and rich question banks could be built quickly. This paper describes the construction of a spreadsheet to generate Calculated questions, using all mathematical functions, and the impact this power tool had on the organization of a course. It catered for large groups by enabling the creation of lots of quizzes to snack from, unlimited rehearsing and immediate feedback of students, allowing lecturers to tutor on demand and students to peer teach. The tool was developed and piloted on a course, "Introduction to Statistics", that was redesigned in order to engage students, improve success rates and minimize redundant lectures. Lectures were reassigned as workshops in which students were challenged with quizzes to activate their learning. Students were invited to collaborate and peer teach. The tool empowered the lecturers to quickly design large question banks. Students approved the self-paced workshops. Confronted with tons of questions in various quizzes students became involved and began to question each other and the lecturers for a way to solve the presented problems.

Keywords

Mathematical functions, spreadsheet, rehearsing, question banks.

Introduction

Teaching large and heterogeneous groups of first year students in a mathematical course, like "Introduction To Statistics" is quite a challenge. Some of the students neglect 'this boring stuff' until the examination, some don't have a clue what it's all about, others find the matter inspiring and there are even experts for which the course holds practically no mysteries.

The old fashioned way of lecturing large groups or smaller parallel groups when large auditoria are unavailable is the opposite of activation. Students are supposed to sit still, listen, ask questions and focus on presented exercises.

To activate these students and have them interacting with each other and the learning content the lecturers wanted to create a workshop environment in which students can work, collaborate and peer teach. The following e-learning scenario was created in order to entice students into such a learning activity. Several question banks were generated and published in small quizzes with one question randomly chosen from a certain category. The students were invited to try these quizzes, preferably in the classroom so they were able to discuss face to face with each other and the lecturers the problems they encountered.

Since statistical functions are not available when building Calculated questions in Moodle, a Learning Management System (LMS), a spreadsheet was developed to overcome this problem. Calculated questions were chosen as the question type as they allow for the creation of multiple questions using wild cards and datasets with numbers.

The following text describes how this spreadsheet was constructed and what impact it had on the course format.

Context

"Introduction to Statistics" is a compulsory, one semester long, course for all first year students in Business Management , a bachelor programme. The course content ranges from frequency tables, calculating mean and

standard deviation to chance theory and distributions like the normal distribution, the binomial distribution and the Poisson distribution. A book with theory and exercises is used during lectures in which lecturers and students walk through the content and take time to solve some of the exercises on paper using calculators. The lectures have a duration of about 120 minutes, take place in classrooms for groups of 30 students and are scheduled as four parallel groups. During examinations students need to solve exercises using their book and notes, pen, paper and a calculator.

Usually there are groups of around 200 students with various mathematical and statistical skills. To counter the heterogeneity of the groups the starting pace of this course is slow, gradually building up. For students with high mathematical skills this is tiresome. This speed is comfortable for students with weaker mathematical skills, but these students need to keep up their learning in order to accelerate. As most of the students have no warm heart for mathematics or statistics arguments are easily found to skip class and refrain from working through the theory and the exercises until the end of the course.

The rationale for the refurbishment of this course is as follows.

- The course format needs to activate students, to get them involved. Students must be able to collaborate and peer teach, live in a spacious computer room and use online forums.
- The course needs to allow for self-pacing, enabling faster students to advance and help slower students.
- The use of spreadsheets should be introduced to solve exercises. As working with spreadsheets has become a critical skill, students need to practice using this tool.
- The course needs to detach from the book so students are more inclined to gather learning content from the web.
- A first line electronic tutor needs to be designed freeing up time for the lecturers to enable them to tutor.

The electronic tutor took on the shape of an electronic course with a large collection of exercises embedded in an LMS. The exercises need to provide the student immediate feedback. Introducing workshops in sequence to the lectures, students get time to work and collaborate in a computer room using the spreadsheet software and the LMS. As students receive feedback from the LMs or each other, the lecturers are able to walk through the room and engage in supportive learning conversations.

Construction

The challenge was to generate a lot of questions given limited time, limited resources and limited functionality in the LMS used by the university college (Blackboard at the time). Given these constraints we choose to use Moodle, as it is open source, free to install and has way more features considering mathematics (Latex filter), question types and feedback settings in quizzes compared with Blackboard.

Calculated Questions

Most LMS incorporate functionality to build questions to be used in tests or quizzes (tests in Moodle are called quizzes). Some LMS's use a question type in which one can define variables. One can then specify a range of values to be used for each variable.

This is an easy example: What is the sum of {a} and {b}?

In this example {a} and {b} are variables, indicated by the curly brackets { }. For these two variables one must then specify a range (from 10 to 100, for example) from which a random number is drawn. The right answer to the question is the number that fits the formula {a}+{b} where {a} and {b} are any number between 10 and 100. Built in Moodle as a Calculated question the result is a question with variables, a set of values for the variables and one or more formulas.

The question type Calculated question was chosen over the alternative type Numerical question. The Numerical question accepts only numerical answers and cannot handle variables. A Calculated question appears in the Moodle question bank as one question with a dataset embedded representing a multitude of questions whereas

the use of Numerical questions would result in several appearances of the same question. When a question needs to be altered only one question needs to be edited when using the Calculated question. In case of the Numerical questions all instances of the same question need to be changed.

Spreadsheets

Moodle allows the creation of Calculated questions based on formulas with variables, but building this kind of questions using the Moodle editor is tedious. There is just a very small edit field with no syntax control. Furthermore, only a limited set of mathematical functions is available in Moodle. As a result, creating statistical formulas like mean, standard deviation, binomial, Poisson and normal distribution is very difficult if not impossible.

A more suitable piece of software for these kinds of calculations is spreadsheets.

However, in order to be able to use spreadsheets for the creation of Calculated questions, the LMS must be capable to understand a file format produced by the spreadsheet program.

Moodle can import and export questions in Moodle XML.

Moodle XML

Extensible Markup Language (XML) is a markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable. Moodle XML is XML with markup declarations defined for Moodle. From an exported set of Calculated questions, created in Moodle, one can easily analyze the structure of this language and create a collection of rules to translate content prepared in a spreadsheet to a file that Moodle understands as a Calculated question with variables, a set of values for the variables and one or more formulas.

Trick

Although we now can translate content prepared in a spreadsheet to a file which Moodle is willing to import and publish as a Calculated question, we still have to find a way to be able to use the rich mathematical functions available in the spreadsheet. For this we need a trick.

For example we would like to use this function, available in a spreadsheet: BINOMDIST(X; trials; SP; C) which returns the individual term binomial distribution probability. Here “X” is the number of successes in a set of trials, “trials” is the number of independent trials, “SP” is the probability of success on each trial and “C” = 0 calculates the probability of a single event and “C” = 1 calculates the cumulative probability.

To calculate this formula we need to specify some parameters X, trials, SP and C. Since these parameters generate a specific result, they must appear in the question text: “What is the chance of precisely (this implies that C must be 0) {X} successes when you know this problem follows a binomial distribution, the chance of success is {SP} for each trial and there are {trials} trials?”

The trick is to calculate the result of this function in the spreadsheet and add the value to a new variable called {Answer}. By doing so we can push to Moodle the following to create a Calculated question: variables {X}, {trials}, {SP} and {Answer} (we don’t need to push {C} since this is fixed in the question text by saying ‘precisely’), the values for all these variables and a formula ={Answer}.

Hidden

Before the variables, their values, the formula(s), the question text and some specific parameters are processed into an XML file some alterations need to be done.

Moodle searches in the question text for the variables used in the formula. Since our formula is ‘={Answer}’, Moodle expects this variable to be present in the text. This implies that when the question is published the answer to the question is visible.

To hide the answer two methods are used. First the answer is made invisible by using HTML. To hide text within html code, one simply surrounds the text with opening and closing comment tags. The opening comment tag is <!-- and the closing tag is -->. Everything in between is hidden from the user. But, they could still see it if they view the source code of the page. Therefore a second method is used to hide the answer. The answer is randomly placed in a random series of numbers. This way when looking at the source code of a published question one sees a string of numbers without thousand or decimal separators enclosed by comment tags.
If needed one can easily create yet another way to obscure the answer.

Check

With the spreadsheet comes two sets of instructions. One set checks the correctness of the used variables and another set generates an XML file from the content in the spreadsheet.

The XML file is then easily imported in a Moodle question bank.

Results

About 60 different question instances where built using this spreadsheet, each with a dataset of 100 values for the variables, representing 6000 different questions (same question text, different values). Complemented with over 100 instances of other type questions, like Multiple Choice questions, Embedded Answer questions and Numerical questions all questions were divided over 20 categories handling seven statistical topics.

A Moodle course was then built containing mainly quizzes and a few forums and links. Three types of quizzes where built for each topic. A quiz with Multiple Choice questions handling mostly statistical definitions. A second type of quiz contains a sequence of Embedded Answer questions building up the analysis of a specific case. And a third quiz was introduced as a ‘snack’ quiz, containing one Calculated question randomly chosen from a relevant category.

The course was then scheduled for one half of the available time as a lecture and the other half as a workshop in which students were invited to work collaboratively on the computer solving the quizzes.

Since two generations of students were eager to take the challenge of the quizzes and preferred the self-paced work in the Moodle course, the lectures were reduced in time to “on demand” lectures allowing more time for the workshops. Having more time to tutor single or small groups of students the lecturers could quickly sort between the stronger and weaker students. The stronger students were then appointed as experts to help tutoring during the workshops.

As a bonus yet another type of quiz was introduced to allow students to do their examination in the Moodle course with automated evaluation. The questions used for examination did not appear in any other quiz.

Success Rates

From qualitative research, organized by the school, students indicated that they appreciate the new working method. They say that after a few hours of lectures, it is nice to do something different, working and collaborating freely.

A reasonable group of students complained about the reduced time for lectures and the lack of a book (although a list with recommended books was available). For them the course seemed chaotic with no book to hold on to, no fixed lectures, no clear tasks. This group needed some structure and some confirmation. They were given a few extra lectures to get them started.

Lecturers were impressed to see the spontaneous emergence of peer tutoring in each group. This was further encouraged by the appointment of student experts on the topic after a learning conversation with the lecturer.

Too many variables were involved in the change of this course. Course material changed, no books were used during lectures, the amount of lectures was strongly reduced, the course was rescheduled from first to second semester and lecturers were reassigned.

So the rise of success rates from 40% up to 70% is not clearly a reaction to this new working format. Nonetheless, this working format was only realizable through the creation of a multitude of tailored question banks in a short period of time with little resources.

Sharing

A webpage is set up to publish several versions of the used spreadsheet, including examples and tutorials. You are welcome to redistribute and / or modify the spreadsheet.

I would like to be informed on who is using the spreadsheet. That way we can exchange ideas and comments in order to improve the spreadsheet.

Please find the spreadsheet here: https://sites.google.com/site/cq4mgenerator/.

Future Work

The discussed spreadsheet is just a tool for optimizing the creation questions. There is still more work to be done. The tool in itself needs to be upgraded to allow less technical lecturers to use it. Another possible improvement would be the development of a plugin to hide the correct answer.

The hard part is still creating enticing questions and embedding them in a supportive learning environment.

With this pilot as example more lecturers are willing to take on the workshop format instead of mere lecturing and the need for stimulating e-learning scenarios to support these workshops is increasing.

Data mining comes into view. Now that we have data produced by generations of students we need to analyze the effectiveness of the question banks we use.