İstanbul Üniversitesi İşletme Fakültesi Dergisi
Istanbul University Journal of the School of Business Administration
Cilt/Vol:40, Sayı/No:1, 2011, 72-85
ISSN: 1303-1732 – www.ifdergisi.org © 2011

Dedicated to the 40th Anniversary of the Journal

# Using computer algebra system to facilitate teaching and learning of probability in business and science schools

***Guoqi Qian***[1]
*Department of Mathematics and Statistics*
*The University of Melbourne*
*Parkville, Victoria 3010, Australia*

***Abstract***

This paper reviews the practice of using computer algebra system (CAS) for facilitating teaching and learning in a second-year level undergraduate subject "Probability for Statistics'' in the University of Melbourne. Some education theory and principles on university teaching are applied to support the practice. In particular, we have adopted Laurillard [1] technological-media-assisted conversational education framework in our teaching and learning of probability for statistics and econometrics.

***Keywords:*** *Computer Algebra System, University Teaching and Learning, Probability, Statistics*

***İşletme ve fen okullarında olasılık öğretim ve öğreniminin kolaylaştırılması için bilgisayar cebiri sistemini kullanmak***

***Özet***

Bu çalışma, Melbourne Üniversitesi'ndeki "İstatistik İçin Olasılık" konulu ikinci sınıf lisans dersinde öğretim ve öğrenimi kolaylaştırmak amacıyla bigisayar cebiri sisteminin (BCS) kullanımının uygulamasını incelemektedir. Uygulamayı desteklemek amacıyla üniversite öğretimi üzerine birtakım eğitim teorisi ve prensipleri kullanılmıştır. Özellikle, Laurillad'ın [1] teknoloji-medya destekli, etkileşimli eğitim sistemi istatistik ve ekonometride olasılık öğretimi ve öğrenimine adapte edilmiştir.

***Anahtar Sözcükler:*** *Bilgisayar Cebiri Sistemleri, Yükseköğretim, Olasılık, İstatistik*

## *1. Introduction*

The subject "Probability for Statistics" is a second-year level undergraduate subject taught by the Department of Mathematics and Statistics of the University of Melbourne. It was developed as a subject component of the new mathematics and statistics curriculum under the Melbourne Model — a tertiary education model developed and implemented by the University of Melbourne since 2008. The model is similar to the European Bologna Model which adopts a three-year undergraduate program followed by advanced courses taught in a two-year master degree or a three-year doctoral program. The "Probability for Statistics" subject was first introduced by me in 2007 to be given to those undergraduate students who are likely to take intermediate statistics and econometrics subjects in their following undergraduate years. The aims of this subject are to not only provide students with a solid training in probability theory, but also teach them a number of important probabilistic methods and rationale for deriving essential

[1] *g.qian@ms.unimelb.edu.au (G. Qian)*

statistics tools for data analysis. This entails the inclusion of additional topics which usually are not taught in traditional probability subjects at the second year level. The involvement of mathematical analysis in this subject is also more intensified and has a wider scope, which means students potentially could take much more time to fully assimilate the subject material if no measure is taken to account for this.

To address these issues, a computer algebra system (CAS) called Maple is used in lectures and taught in computer lab tutorials of this subject. The underlying pedagogic rationale is that students will then be able to mainly focus on understanding and connecting the concepts and models related to probability distributions and random variables, and to focus on the principles and procedures of the many techniques to be learned. The complicated but routine parts of the various technical derivations can be left to be done by the computer algebra system. This paper will discuss how the use of CAS can facilitate the teaching and learning process of probability and statistics. This discussion is motivated by my current teaching of the subject "Probability for Statistics" and my recent reading of the book by Diana Laurillard [1], the Chapter 4 of Paul Ramsden [2] and a number of other papers on higher education teaching and learning. This paper will also discuss the impact of using CAS on a subject's assessment and curriculum design.

The paper is organised as follows. In section 2 an overview of the subject "Probability for Statistics" is given including the subject description, the logistics motivation and education justification of teaching this subject to students and a brief account on using computer algebra system for facilitating student learning. In section 3 an outline is provided to summarise Laurillard's theory on teaching-learning conversational framework using technological media. This theory is used to validate the practice of using CAS for assisting teaching and learning in "Probability for Statistics". In section 4 explanations will be provided on what is a CAS as a technological and educational medium, why it is used in "Probability for Statistics", how it is designed and organised for teaching and learning in "Probability for Statistics", how it will affect students taking different learning approaches, and how assessment and feedback need to be properly structured to reflect the use of CAS. Remarks will be provided in section 5 to conclude the paper.

## *2. The Subject "Probability for Statistics"*

The subject "Probability for Statistics" is a second-year level undergraduate subject. It has been offered as a key component in the new mathematics and statistics undergraduate program under the Melbourne Model since year 2008. The contents of this subject are attached as an appendix at the end of this paper. The description of the subject can be found from the University of Melbourne Course Handbook (Mathematics and Statistics) as following:

> This subject develops the probability theory that is necessary to understand statistical inference. Properties of probability are reviewed, random variables are introduced, and their properties are developed and illustrated through common univariate probability models. Models for the joint behaviour of random variables are introduced, along with conditional probability and Markov chains. Methods for obtaining the distributions of functions of random variables are considered along with techniques to obtain the exact and approximate distributions of sum of random variables. These methods will be illustrated through some well known normal approximations to discrete distributions and by obtaining the exact and approximate distributions of some commonly used statistics. Computer packages are used for numerical and theoretical calculations but no programming skills are required.

Apart from "Probability for Statistics", the Department of Mathematics and Statistics already has an existent second-year level probability subject "Probability" which shares substantial amount of common description with "Probability for Statistics". So what are the motivation and justification for the department to establish the seemingly redundant subject "Probability for Statistics"?

I was not deeply involved in the development of "Probability for Statistics" at its initialization stage at the management level and had taught "Probability" for only one time before, but I was assigned to develop the course material for the new subject and have been teaching it since then. A reflection from this experience has enabled me to understand that, although the motivation behind developing "Probability for Statisticsa" mainly comes from the logistics considerations, its justification is mainly the learning and teaching benefits.

The subject "Probability" used to have a large enrollment size of approximately 300 students before "Probability for Statistics" was introduced. And the department expected to get many of them to take statistics subjects in their third year studies. Among the enrolled, about the half number were actuarial science students, which was the result of the university's Center for Actuarial Science prescribing "Probability" to their students. The past enrollment statistics suggest that these actuary students are very unlikely to take any mathematics and statistics subjects at the third year level; and really there is little room in their course program for them to do so. The subject contents of "Probability" are negotiated between the Department of Mathematics and Statistics and the Center for Actuarial Science to be conducive to learning by the actuary students who usually have much stronger mathematics background than the business and other sciences students enrolled in this subject. Due to this particular design feature the non-actuarial science students enrolled in the subject "Probability" are likely to encounter more severe learning difficulties than their actuary fellow students, thus their learning outcomes are less likely to meet the expectations set by the subject. It has been found difficult, with regard to teaching and learning in "Probability", to reconcile the disparities between the mathematics backgrounds of the various student cohorts enrolled, to harmonize the diversities between their respective appropriate ways of apprehending the subject knowledge, and to coordinate the differences between the respective disciplines' learning outcome planning. Consequently, we experienced insufficient enrollment in our third year statistics subjects before the subject "Probability for Statistics" was introduced to accompany "Probability". As it became an important logistical goal of the department to increase the third-year statistics subject enrollments, it required the department to take immediate action so that more students became interested in statistics and statistics as a discipline was viable. I believe it was mainly this pressing logistics consideration that led to the department to create the new subject "Probability for Statistics".

The introduction of "Probability for Statistics" brings not only institutional and logistical benefits to the department but also the learning benefits to the targeted student cohort and professional benefits to the statistics discipline. Notwithstanding this, introducing and teaching "Probability for Statistics" entails us to take challenges on designing new relevant teaching materials, setting up the associated learning context and delivering them to achieve the projected outcomes. It is not the intention of this paper to deliberate about the subject content or academic knowledge of this subject. Rather, I will discuss the pedagogical consideration and practice used to facilitate teaching and learning involved in this subject. The focus will be on the use of a computer algebra system called Maple in our teaching and learning practice. The much anticipated higher education monograph "Rethinking University Teaching: a conversational framework for the effective use of learning technologies" [1] provides a sound and compelling education theory for supporting and interpreting the designing and delivering practice we have used in "Probability for Statistics".

## 3. *Effective Teaching and Learning in Conversational Environment Equipped with Interactional Technologies*

Through analysing the teaching and learning processes and devising teaching strategy for successful learning, Laurillard has derived in her 2002 book aforementioned a model she called the conversational framework for attaining effective teaching and quality learning in higher education.

### *3.1. Teaching as Mediating Learning*

Laurillard [1] perceives teaching as a rhetorical activity: it mediates learning, allowing students to acquire academic knowledge — the knowledge of someone else's way of experiencing the world or the so-called second-order experience of the world. She also notes academic learning is different from everyday percepts learning which is done through situated cognition in natural environments. Thus teaching must create artificial environments that afford the academic learning of "precepts", i.e. descriptions of the world [1, p.24].

### *3.2. Issues in Related to Teacher-Students Interaction*

Moreover, teaching needs to address the following aspects of what students bring with them to learning a new topic [1, p.40]:

- conceptions of the topic — teachers need descriptions of the ways students conceptualise a topic to be able to challenge their fundamental misconceptions;
- representational skills — students need explicit practice in the representation of knowledge of their subject, in language, symbols, graphs and diagrams, and in the manipulation and interpretation of those representations;
- an epistemology — teachers must enable students to develop their epistemological and ethical beliefs, and in particular, their conceptions of learning.

### *3.3. Interdependent Aspects of the Learning Process*

On the other hand, investigation of successful learning activities has revealed five interdependent aspects of the learning process [1, pp.60-61]:

- apprehend the structure of the discourse — e.g. focus on the narrative line, distinguish evidence and argument, organise and structure the content into a coherent whole;
- interpret the forms of representation — e.g. practise mapping between the concept, system, event or situation and its representation, practise using the forms of representation of an idea, represent the discourse as a whole as well as its constituent parts;
- act on descriptions of the world — e.g. combine descriptions and representations to generate further descriptions of the world, manipulate the various forms of representation of the world;
- use feedback — e.g. use both intrinsic and extrinsic feedback to adjust actions to fit the task goal, and adjust descriptions to fit the topic goal;
- reflect on the goal-action-feedback cycle — e.g. relate the feedback to the goal or message of the discourse, reflect on how the link between action and feedback relates to the structure of the whole.

### 3.4. Deep Holistic Approach versus Surface Atomistic Approach in Learning

Investigation of students' learning activities has also identified two contrasting approaches to acquiring academic knowledge from a subject. One is known as the "deep approach", where a student looks for the meaning and processes the content in a "holistic" way, preserving the original structure of the discourse and therefore preserving its intended meaning. The other is known as the "surface approach", where the student focuses on key words or phrases and processes the text in an "atomistic" way, distorting the original structure and therefore changing its meaning. Marton and Säljö [3,4] and Svensson [5] gave the earliest studies of these two approaches. A more recent study can be found in Ramsden [2].

### 3.5. Teaching strategy

Given what are known about the characteristics of teaching and student learning, Laurillard has argued for a best teaching strategy, which she expresses as an iterative dialogue between teacher and student focused on a topic goal. The responsibilities of both teacher and student can be grouped as four distinct aspects of the progression of the dialogue [1, pp.77-78]:

**Discursive**

- Teacher's and student's conceptions should each be continually accessible to the other;
- teacher and student must agree on learning goals for the topic;
- the teacher must provide a discussion environment for the topic goal, within which students can generate and receive feedback on descriptions appropriate to the topic goal.

**Adaptive**

- the teacher has the responsibility to use the relationship between their own and the student's conception to determine the task focus of the continuing dialogue;
- the student has the responsibility to use the feedback from their work on the task and relate it to their conception.

**Interactive**

- the teacher must provide a task environment within which students can act on, generate and receive feedback on actions appropriate to the task goal;
- the student must act to achieve the task goal;
- the teacher must provide meaningful intrinsic feedback on their actions that relates to the nature of the task goal.

**Reflective**

- the teacher must support the process in which students link the feedback on their actions to the topic goal for every level of description within the topic structure;
- the student must reflect on the task goal, their action on it, and the feedback they received, and link this to their description of their conception of topic goal.

### 3.6. Conversational Framework and Forms of Educational Media

A conversational framework for the learning process is thus formed based on the teaching strategy with the aforementioned four aspects. This framework in turn forms a basis for analysing the main technology-based educational media in terms of to what extent they support effective teaching and quality learning. Table 1 is from Laurillard [2] and summarises the five principal media forms with the learning experiences they support and the methods used to deliver them. The computer algebra system that we use in the subject "Probability for Statistics" can be regarded as a technology-based educational medium; and it possesses attributes of three media forms mentioned in Table 1: interactive, adaptive and productive.

**Table 1 Five Principal Media Forms, Learning Experiences and Delivering Methods**

| *Media forms* | *Learning experience* | *Methods/technologies* |
|---|---|---|
| Narrative | Attending, apprehending | Print, TV, viedo, DVD |
| Interactive | Investigating, exploring | Library, CD, DVD, Web resources |
| Communicative | Discussing, debating | Seminar, online conference |
| Adaptive | Experimenting, practicing | Laboratory, field trip, simulation |
| Productive | Articulating, expressing | Essay, product, animation, model |

(Source: D. Laurillard [1, p.90])

## *4. Computer Algebra System as a Technology-Based Educational Medium*

### *4.1. What is a Computer Algebra System?*

A computer algebra system (CAS) is a software program that facilitates manipulation of mathematical formulae. The primary goal of a computer algebra system is to automate tedious and sometimes difficult algebraic manipulation tasks. The principal difference between a computer algebra system and a traditional calculator is the ability to deal with mathematical expressions symbolically rather than numerically. For example, a computer algebra system can calculate the solution of a quadratic equation $ax^2+bx+c=0$ to be $x=(2a)^{-1}(-b\pm\sqrt{b^2-4ac})$ which is a mathematical expression in terms of symbols *a*, *b* and *c*, while a traditional calculator cannot compute the solution unless *a*, *b*, *c* are assigned some numerical values. A computer algebra system can do many types of symbolic mathematics including factorization, taking some limits, differentiation, some integration, summation of infinite series, solving equations, transformations and so on. Computer algebra systems often include facilities for typesetting and graphing mathematical expressions, and provide a programming language for the user to define his/her own procedures. Some computer algebra systems further include toolboxes for specific areas of applications such as probability, statistics and physics. A current leading commercial computer algebra system is Maple which is the one we use in the subject "Probability for Statistics". A Maple console window is shown in Figure 1.

Computer algebra systems have not only changed how mathematics is taught in many universities, they also provide a very effective tool for teaching and learning other subjects which use mathematics results but do not particularly concern how these results are derived. To illustrate the effectiveness, we provide two examples below.

In the first example, suppose in a probability lecture, the teacher needs the result $1+2^{-2}+3^{-2}+\cdots=\sum_{k=1}^{\infty}k^{-2}=6^{-1}\pi^2$ for calculating a discrete probability. The teacher would explain in full detail to students on how this result is obtained if it were a calculus lecture. In the probability lecture, the teacher's focus is on describing the discrete probability concept in this case, thus should not spend too much time and resource to get into the detail of calculating $\sum_{k=1}^{\infty}k^{-2}$. So Maple would be very useful here. Typing the command "`sum(1/k^2,k=1..infinity)`" in Maple would give the desired result. In addition, the teacher can easily use Maple to compute $1+2^{-2}+\cdots+100^{-2}$, $1+2^{-2}+\cdots+1000^{-2}$ and $1+2^{-2}+\cdots+10000^{-2}$ etc. to assist students to understand that $\sum_{k=1}^{\infty}k^{-2}=6^{-1}\pi^2$ is true.

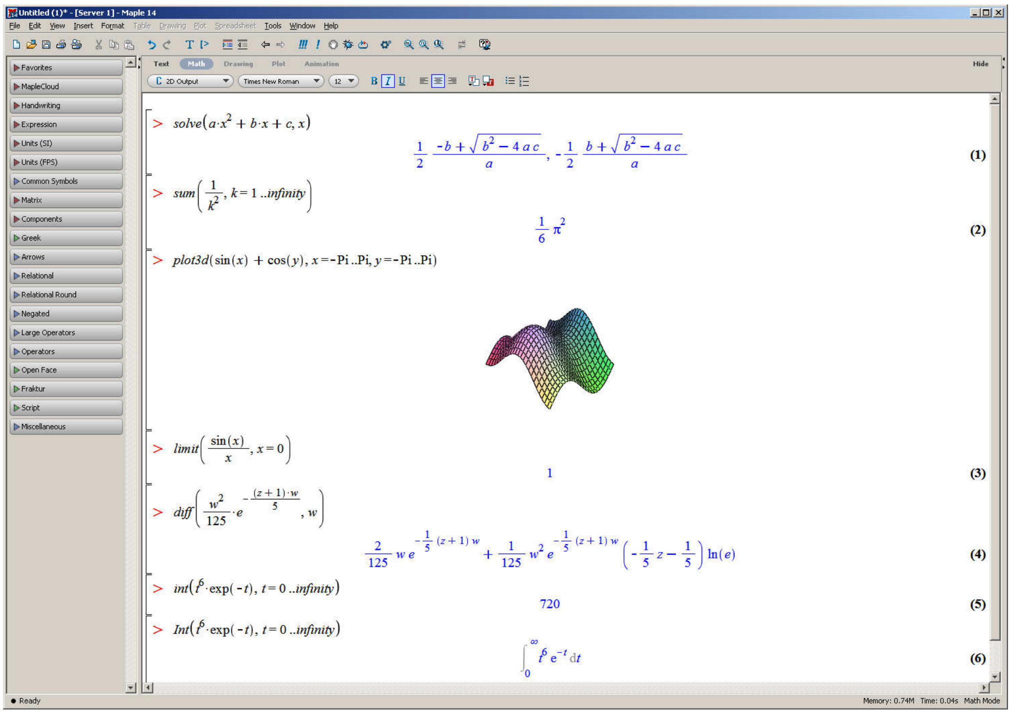


**Figure 1 Maple 14 Console Window**

In the second example, suppose students have an assignment task related to gamma distribution (a common probability distribution for describing amount of change, e.g. amount of monthly rainfalls in millimeters, amount of yearly insurance claims in dollars, etc.), in which they typically need to calculate an integral like $\int_0^\infty t^6 e^{-t} dt$. Using Maple with command "`int(t^6*exp(-6), t=0..infinity)`" this integral can be found to be 6!=720. Without a computer algebra system a typical second year student will need to use the integration-by-parts technique to find the result which would be very tedious and likely to incur errors.

## 4.2. *Why CAS is Used in "Probability for Statistics" Teaching and Learning?*

The rationale of using a computer algebra system in "Probability for Statistics" is both logistical and pedagogical. As seen in Section 2 and Appendix, the subject covers a wide range of topics including basic probability models, various types of random variables and distributions, their properties, and their applications in statistics. Describing, reasoning and apprehending these topics will directly involve many mathematics ideas, techniques and results at the level of the first-year undergraduate calculus and algebra subjects. Given that there are only 36 lecture hours for teaching this subject, it is not affordable to spend too much time and context on tediously practising the mathematical manipulations needed for imparting the probability knowledge to students. But if these mathematical manipulations are not performed in teaching, conveying those probability ideas, properties and principles to students is likely to be like building a castle on sand, which will fall apart very soon in students' minds. Thus teaching probability without involving mathematics manipulations will not achieve its targeted goal and not result in quality learning by students. Therefore, it would be very helpful to find a technology-based educational medium to be used for facilitating the required mathematical manipulations, so that teaching and learning probability can still be built on the necessary mathematics background but will focus on imparting and apprehending probability knowledge. It becomes clear that the computer algebra system Maple can fulfill this task.

Recall in Table 1 in Section 3.6 five principal media forms used in teaching and learning are summarised together with the learning experiences they support. In comparison with these five principal media forms, we can claim that Maple possesses all the attributes of interactive, adaptive and productive media. In other words, Maple can support the learning experiences of investigating, exploring, experimenting, practising, articulating and expressing in the context of teaching and learning probability for statistics. We provide in the following subsection a case study used in "Probability for Statistics" to support our claim.

### *4.3. A Case Study on Learning Central Limit Theorem*

Central Limit Theorem (CLT) is a fundamental probability property in a heutistic form saying that, when the sample size goes to infinity and some regularity conditions are satisfied, the probability distribution of the standardized average of any random sample converges to the standard normal distribution $N(0,1)$. The normal distribution $N(\mu, \sigma^2)$ usually can be represented by a bell-shaped symmetric probability density curve. Specifically, the normal probability density curve is given by the function $f(t;\mu,\sigma) = (2\pi\sigma^2)^{-1/2}\exp\{-(2\sigma^2)^{-1}(t-\mu)^2\}, \quad -\infty < t < \infty,$ where μ is the mean parameter and $\sigma$ is the standard deviation parameter. So the CLT can be equivalently stated as the probability density curve of the standardized average of a random sample converges to the standard normal density curve $f(t; 0,1)$ when the sample size goes to infinity. The effects and applications of CLT can be seen in many real world practices. Note that the statement in CLT is quite abstract because it is expressed for "any" random sample and the sample size goes to "infinity".

In "Probability for Statistics", students are expected to understand the CLT and further be able to prove it using the calculus method and some new concepts (i.e. moment generating function or characteristic function) together with their properties which they will have already learned by the time of learning CLT. To describe the CLT in a form that is concrete and that students can visualise, a random variable $X_1$ with its distribution is specified (suppose $X_1$ has mean 0 and standard deviation 1 for convenience of presentation). In particular, we specify the probability density function of $X_1$ to be $f(x) = \sqrt{243/500}x^2, \quad -\sqrt{5/3} < x < \sqrt{5/3}$. Then by Maple the probability density curve of $X_1$ is graphed and its various attributes are calculated and listed. The density curve and the characteristic function for $X_1$ are compared with those of the standard normal random variable. Students will clearly see they are very different. The next step is to define a second random variable $X_2$ which has the same distribution as $X_1$ but is independent of $X_1$. Then compare the density curve and the characteristic function of the standardized $W_2 = (X_1 + X_2)/\sqrt{2}$ with those of the standard normal random variable (See Figure 2). Students now see that the difference is smaller than before. The third step is to compare the density curve and characteristic function of $W_3 = (X_1 + X_2 + X_3)/\sqrt{3}$ with those of the standard normal random variable. After a few more steps students will be able to imagine that the distribution of $W_n = (X_1 + X_2 + \cdots + X_n)/\sqrt{n}$ will eventually converge to the standard normal distribution (See Figure 3 for $n$=3 and 8, and compare it with Figure 2 for $n$=2). Thus they will have an initial understanding of CLT. Now it is their turn to repeat the same exploration but for $X_1$ $X_2$, …, $X_n$, … having a different distribution. They do this in their computer lab class or homework through interaction with Maple. They may not be able to succeed with just one try. But they do not have to wait for the lecturer's feedback to continue trying, because Maple is able to give them instant feedback. So they can adapt their approach and keep on trying until success. The proof of CLT can be done in a similar way: first use Maple to do the proof for specified cases in a goal-action-feedback style to ensure students' understanding of the procedure; then do the proof by hand for the general case after articulating, relating and digesting all that have been learned so far.

Through using Maple for teaching the CLT and more generally many other topics in "Probability for Statistics", a favourable environment has been created for encouraging students to use a "deep holistic" approach to achieve their learning objectives. Using a "surface atomistic" approach works only when a student's goal is repeating the learned examples; it will not get through Maple's feedback system if the student wants to fully understand the subject topics and to be able to apply them for practising. Maple will provide instant intrinsic feedbacks and continue to do so until the student takes all feedbacks, makes directed changes and eventually succeeds in applications.

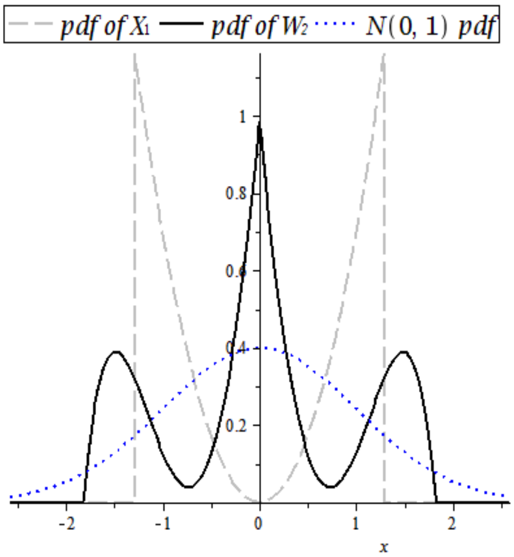


**Figure 2 Comparison of PDF Curves of $X_1$, $W_2$ and Standard Normal**

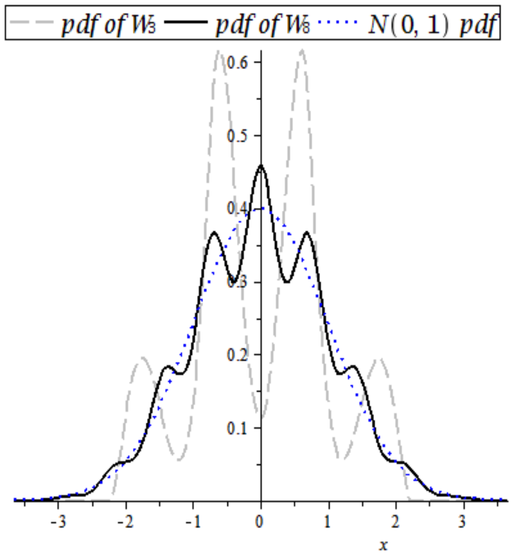


**Figure 3 Comparison of PDF Curves of $W_3$, $W_8$, and Standard Normal**

### *4.4. Organising Maple in "Probability for Statistics" Teaching and Learning*

Students need to learn the computer algebra system Maple before they can use it for helping them learn probability. But Maple is very user-friendly, and operating Maple is just like operating an advanced graphic calculator. In "Probability for Statistics" we use the first computer lab class for training students to use Maple; and this training seems being sufficient for students to be able to use them largely independently for analysing standard questions. Then in each of the following weekly computer lab classes students will be tutored to apply Maple to progressively assist their learning of more and more complicated topics taught in lectures or demonstrated in practice classes. They are further asked to use Maple to do some assignment questions to articulate the knowledge and skills that they have learned. A sample of the computer lab questions and solutions, which were used in "Probability for Statistics" in semester 1 of 2010, are provided in Section 6.2 for illustration.

### *4.5. Effects of Maple on Assessment of Learning and Feedback to Teaching*

The assessment of learning must be aligned to reflect the use of the computer algebra system Maple in teaching and for practising, if Maple is so used. In "Probability for Statistics" a mid-semester computer based test worth 10% of the total assessment is designed and executed to assess students' learning outcomes with respect to using Maple for studying probability. A sample of the computer based test used in Semester 1 of 2010 is given in Section 6.3 for reference. In the computer based test, students are not required to do programming and are not tested on their ability of memorising commands either; they are tested on the ability of using Maple to solve probability problems. The mid-semester test is able to provide useful feedbacks to the lecturer so that he can respond and still have time to implement changes, if necessary, into the remaining teaching of the subject. Students also have the opportunities to apply Maple to solve some of their assignment questions. In total, the Maple involved assessment is about 20% of the total assessment.

Because the subject assessment contains an element accounting for the role of Maple, students would know that using Maple is relevant to their learning the subject. But they will gradually realise that they do not need to learn the full content of Maple for being comfortable to use it. The necessary knowledge that they need to know is defined and provided in the computer lab material progressively. The solutions of the computer lab questions and assignments are provided to students as well after each lab class through the subject web page. Further, tutors in the computer labs are requested to provide their comments and collect students' feedbacks after each lab class, to see whether any adjustment is needed for the subsequent labs and to implement it if so.

## *5. Concluding Remarks*

In this paper I give a reflection on the teaching and learning of the subject "Probability for Statistics'' which I taught as a second-year subject over the last four years in the University of Melbourne. The reflection is mainly on the role of the computer algebra system Maple in teaching and learning. It is the outcome of a happy combination of my teaching experience from this subject and my reading the book by Laurillard [1] and other references on higher education. I believe a continuous reflection process during teaching and learning is always necessary for achieving the best learning outcomes for students. It certainly has contributed to my improving teaching and learning practice of "Probability for Statistics" and other subjects in general.

## *References*

[1] D. Laurillard, *Rethinking University Teaching, A Conversational Framework for the Effective Use of Learning Technologies (2nd edition),* London: RoutledgeFalmer, (2002).

[2] P. Ramsden, *Learning to Teach in Higher Education, (2nd edition),* London: RoutledgeFalmer (2003).

[3] F. Marton and R. Säljö, On qualitative differences in learning I: outcome and process. *British Journal of Eduactional Psychology*, 46, 4-11 (1976a).

[4] F. Marton and R. Säljö, On qualitative differences in learning II: outcome as a function of the learner's conception of the task. *British Journal of Eduactional Psychology*, 46, 115-127 (1976b).

[5] L. Svensson, On qualitative differences in learning III: study skills and learning. *British Journal of Eduactional Psychology*, 47, 233-243 (1977).

## *Appendix*

- ***Subject Contents of "Probability for Statistics"***

*Module 1. Probability*

§1.1. Introduction to probability

§1.2. Definitions and properties related to probability

§1.3. Classic probability models

§1.4. Conditional probability and multiplication rule

§1.5. Independent events

§1.6. Law of total probability, Bayes' theorem, prior and posterior probabilities

*Module 2. Discrete Distributions*

§2.1. Discrete random variables

§2.2. Mathematical expectation

§2.3. Mean, variance and standard deviation

§2.4. Bernoulli trials and the binomial distribution

§2.5. The moment-generating function

§2.6. The Poisson distribution

*Module 3. Continuous Distributions*

§3.1. Continuous random variables

§3.2. The uniform and exponential distributions

§3.3. The gamma and chi-square distributions

§3.4. Distributions of functions of a continuous random variable

*Module 4. Multivariate Distributions*

§4.1. Distributions of two random variables

§4.2. The correlation coefficient

§4.3. Conditional distributions

§4.4. Multivariate transformations of random variables

§4.5. Several independent random variables

§4.6. Distributions of sums of independent random variables

§4.7. Chebyshev's inequality and convergence in probability

*Module 5. The Normal Distribution*

§5.1. The normal distribution

§5.2. Random functions associated with normal distributions

§5.3. The central limit theorem

§5.4. Approximations for discrete distributions

§5.5. Limiting moment-generating functions

§5.6. The bivariate normal distribution

*Module 6. Introduction to Stochastic Processes and Markov Chains*

- ***"Probability for Statistics" Week 7 CAS Lab Questions and Solutions, Semester 1 2010***

In this lab, we will study continuous random variables and their distributions using Maple.

Let *X* be a continuous random variable with probability density function (pdf) *f*(*x*). A *typical Maple session* for studying *X* and its distribution is given below:

1. `> with(Statistics):` #This command calls for the package `Statistics`.
   `> assume(...);`
   This command is necessary only if you want to put some restrictions on the domain of some parameters. E.g. `assume(s>0, t<=0)`.
2. Create the pdf function *f*(*x*) if *X* does not have a built-in distribution object in Maple:
   `> f:="expression of the function";    > f:=unapply(f,x);`
   (Type `> help(Distributions);` to see what built-in distributions are available.
   Also note that `f:=unapply(f,x)` turns the expression `f` into a functional operator so that you can now use `f(b)` to evaluate `f` at `x=b` instead of using `subs(x=b,f)`. )
3. Create the random variable *X*:
   `> X:=RandomVariable(Distribution(PDF=f));` or
   `> X:=RandomVariable("Name and attributes of the distribution");`
   if the distribution of *X* is a built-in one in Maple.
4. Recall the pdf *f*(*x*) of *X* or calculate *f*(3):
   `> PDF(X,x);` or `> PDF(X,3);`
5. Call the cumulative distribution function (cdf) of *X* or evaluate the cdf at `x=3`:
   `> CDF(X,x);` or `> CDF(X,3);`
6. Calculate the probability *P*(*X*>2) and *P*(1≤ *X*<2):
   `> Probability(X>2);` and `> Probability({X>=1, X<2});`
7. Find the moment generating function (mgf) of *X*:
   `> MGF(X,t);` or `> int(exp(t*x)*PDF(X,x),x=-infinity..infinity);`
8. Find the mean of *X*:
   `> Mean(X);` or `> Moment(X,1);` or `> int(x*PDF(X,x),x=-infinity..infinity);`
9. Find the variance of *X*:
   `> Variance(X);` or `> Moment(X,2)-(Moment(X,1))^2;`
10. Find, say 95th percentile of *X*:
    `> Percentile(X,95);` or `> solve(CDF(X,x)=0.95,x);`
11. Generate a sample of 20 observations from the distribution of *X* and plot the histogram:
    `>A:= Sample(X,20); > Histogram(A,range=a..b);`
12. Plot the density of *X*:
    `> plot(PDF(X,x),x=a..b);` or `>DensityPlot(X,range=a..b);`

**Questions:**

1. Suppose a value is chosen "at random" in the interval [0, 8]. In other words, *x* is an observed value of a random variable $X \sim U(0,8)$. The random variable *X* divides the interval [0, 8] into two subintervals, the lengths of which are *X* and 8−*X* respectively. Denote by $Y = \min(X, 8-X)$, the length of the shorter one of the two intervals. Find the cdf of *Y*. Also what is the name of the distribution of *Y*?
2. Let random variable *X* have the pdf $f(x) = e^{-x}(1+e^{-x})^{-2}, \quad -\infty < x < \infty$.
   a) Find the cdf of *X*.
   b) Find the mean and variance of *X*.
   c) Find $P(3 < X < 5)$.
   d) Find the 85th percentile of *X*.
   e) Let $Y = (1+e^{-X})^{-1}$. Find the cdf of *Y*. Can you tell the name of the distribution of *Y*?
3. Let *X* have an Exponential distribution with mean *θ*. Assume *θ* >0, *u*>0 and *v*>0.
   a) Find *P*(*X*>*u*).
   b) Find *P*(*X*>*u*+*v*).
   c) Find *P*(*X*>*u*+*v*|*X*>*v*).

d) Compare results of (a) and (c). What conclusion can you draw? What is the reason behind the conclusion?

4. (*This question is to illustrate the relationship among Gamma, Chisquare and Poisson distributions*.) Let *X* be a Gamma random variable with mean parameter *θ*=2 and shape parameter $\alpha = 8$. That is, the pdf of *X* is $f(x) = (2^8\Gamma(8))^{-1}x^{-7}e^{-x/2} = 1290240^{-1}x^{-7}e^{-x/2}$, $x \ge 0$.
   a) Use Maple to find the mgf of *X* and calculate *P*(X>*3*).
   b) Create a Poisson random variable *Y* with parameter $\lambda = 3/2$. Then compute the probability *P*(*Y*≥*7*). Compare the result with *P*(*X*>3) in (a). Why they are the same?
   c) Create a ChiSquare random variable *Z* with degrees of freedom 16. Find the mgf of *Z* and *P*(*Z*>3). Are they the same as the mgf of *X* and *P*(*X*>3)? Why?

**Solutions:**

1.
```
> with(Statistics):
> X:=RandomVariable(Uniform(0, 8)); > Y:=min(X,8-X); > CDF(Y,y);
```
*Y* has a Uniform(0,4) distribution.

2.
```
> with(Statistics):
a) > f:=exp(-x)/(1+exp(-x))^2; > f:=unapply(f,x);
   > X:=RandomVariable(Distribution(PDF=f));
   > CDF(X,x); or > int(PDF(X,x),x); or > int(f(x),x);
b) > Mean(X); > Variance(X);
c) > Probability({X>3, X<5}); or > int(PDF(X,x),x=3..5); or > int(f(x),x=3..5);
d) > Percentile(X,85); or > solve(CDF(X,x)=0.85,x);
e) > Y:=1/(1+exp(-X)); > CDF(Y,y); # Y has a Uniform(0,1) distribution.
```

3.
```
> with(Statistics):
> X:=RandomVariable(Exponential(theta)); > assume(theta>0, u>0, v>0);
> Probability(X>u); > Probability(X>u+v); > Probability(X>u+v)/Probability(X>v);
> simplify(%);
```
One can conclude that *P*(*X*>*u*+*v*|*X*>*v*)=*P*(*X*>*u*). This is true because of the "no-memory" property of the exponential distribution.

4.
```
> with(Statistics):
a) > X:=RandomVariable(Gamma(2,8)); > assume(t<1/2);
   > MGF(X,t); or > int(exp(t*x)*PDF(X,x),x=0..infinity);
   > Probability(X>3); or > int(PDF(X,x),x=3..infinity);
b) > Y:=RandomVariable(Poisson(3/2));
   > Probability(Y<=7); or > sum((3/2)^k*exp(-3/2)/k!,k=0..7);
```
Same as *P*(*X*>3). The event that `waiting time' for the 8th change is greater than 3 time units is the same as the event that there are at most 7 changes during 3 units of time.
```
c) > Z:=RandomVariable(ChiSquare(16));
   > MGF(Z,t); > Probability(Z>3);
```
*Z* and *X* have the same mgf, distribution and probability. ChiSquare distribution with df=*r* is defined as a Gamma distribution with *θ* =2 and $\alpha = r/2$.

- ***"Probability for Statistics" Computer Based Test Questions and Answers, Semester 1 2010***

The duration of the test is 50 minutes. In the test you are permitted to use Maple and any other notes to answer your questions. You may ask your tutor for help with technical difficulties on using Maple and the computer. You must hand this sheet back to your tutor before you leave the lab. You must not communicate with other students during the test; and do not talk about this test with any other students until the overall testing period is complete. The marks for each question are indicated in brackets [ ]. The total marks are 40.

**Question 1.** [12marks]

(*Banach Matchbox Problem*) A cigarette smoker carries two matchboxes, one in his right pocket and one in his left pocket. Whenever he wants to smoke, he selects a pocket at random and takes a match from the box in that pocket. Suppose each box contains 100 matches initially. Let *X* be the number of matches left in a box when the smoker for the first time discovers that the other box is empty. It can be shown that the pmf of *X* is $P(X=x)=\binom{200-x}{100}2^{x-200}, \quad x=0,1,2,\cdots,100$.

Complete the following tasks and keep 4 significant digits after the decimal point in your answers.

a) Find the probability $P(X \geq 10)$.
b) Find the mean $E(X)$.
c) Find $E(X^2)$.
d) Find $E((X+1)^{-2})$.

**Question 2.** [14marks]

Let a continuous random variable *X* have the pdf $f(x)=\frac{2}{9}(x+1)(2-x), \quad -1<x<2$.

a) Find the cdf *F*(*x*) of *X*.
b) Find the probability *P*(-3.5<*X*<1.8).
c) Find the mean $E(X)$.
d) Find the mgf $M(t)=E(e^{tX})$.
e) Find the third moment $E(X^3)$.
f) Let $Y=X^2$. Find the support of *Y*. Also find the pdf *g*(*y*) of *Y*.

**Question 3.** [14marks]

Consider random variables *X* and *Y* which have the joint pdf $f(x,y)=\frac{6}{7}(x+y)^2, \quad 0<x<1, 0<y<1$.

a) Sketch a graph of the support of *X* and *Y*.
b) Find the marginal pdf $f_1(x)$ of *X*.
c) Find the mean *E*(*X*).
d) Find the variance *Var*(*X*).
e) Find the covariance *Cov*(*X*,*Y*).
f) Find the correlation coefficient *ρ* between *X* and *Y*.
g) Find the probability *P*(2*X*+*Y*>1).

**Answers:**

**Question 1.** a) 0.4694; b) 10.3260; c) 169.0219; d) 0.08862.

**Question 2.** a) $F(x)=0 \text{ if } x \leq -1; \ =\frac{1}{27}(7-2x)(x+1)^2 \text{ if } -1<x\leq 2; \text{ and} =1 \text{ if } x>2$; b) 0.9873; c) $\frac{1}{2}$;

d) $M(t)=\frac{2}{9t^3}(3e^{-t}t+2e^{-t}+3e^{2t}t-2e^{2t})$; e) $\frac{4}{5}$;

f) $g(y)=\frac{4}{9\sqrt{y}}-\frac{2\sqrt{y}}{9} \text{ if } 0\leq y<1; \ =\frac{(\sqrt{y}+1)(2-\sqrt{y})}{9\sqrt{y}} \text{ if } 1\leq y<4; \text{ and} =0 \text{ elsewhere}$.

**Question 3.** a) omit. b) $f_1(x)=\frac{2}{7}+\frac{6}{7}(x+x^2), 0<x<1$; c) $\frac{9}{14}$; d) $\frac{199}{2940}$; e) $\frac{-5}{588}$; f) $\frac{-25}{199}$; g) $\frac{15}{16}$