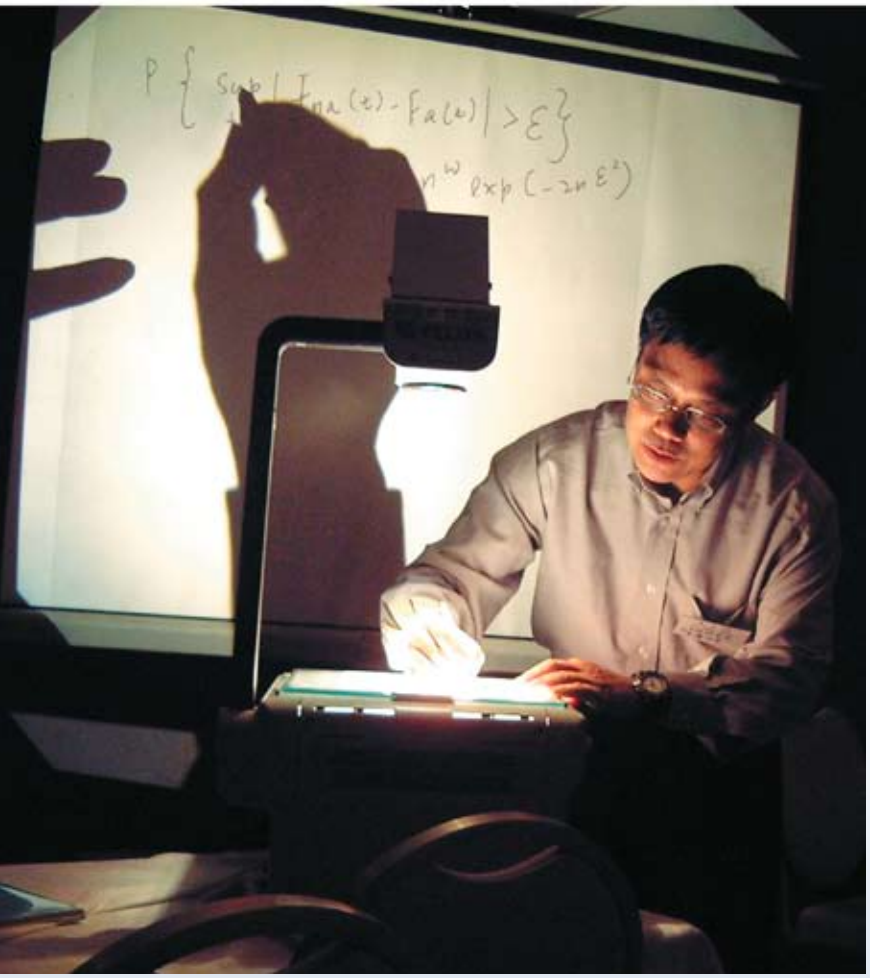
A Distinguished Figure in Statistics
統計學人分享統計與生活



In July, Professor Zhu Lixing of HKBU's Department of Mathematics added to the many honours he has received when he was one of the few non Americans to be selected as Fellow of the American Statistical Association in 2007. The prestigious award was for his outstanding contributions in several areas of statistical research. These areas sound highly technical but in fact have important applications in many areas of our life.

今年七月，屢獲殊榮的浸大數學系朱力行教授，憑藉他對統計學的卓越貢獻，獲選為二零零七年度美國統計協會院士，成為獲得這項殊榮的少數非美籍人士。朱教授的研究涉及的知識非常專門，但卻與我們的日常生活息息相關。

Understandably, Professor Zhu is proud of his award from the American Statistical Association, which is the largest academic body concerned with statistics. Professor Zhu is the only Asian specialist to be honoured with the Humboldt Research Award granted by the Alexander von Humboldt Foundation of Germany. In addition, he was selected as an Institute of Mathematical Statistics Fellow, and elected Director of International Chinese Statistical Association.

Professor Zhu was honoured for, among other things, his work in dimension reduction methods, goodness-of-fit testing, and empirical likelihood – not subjects that many people can say they are familiar with.

But in fact these research areas have multiple practical uses, particularly in

美國統計協會是全球最大的統計學學術組織，朱教授能夠得到他們認同，當然深感自豪。另外，他也獲美國數理統計學院頒發院士、國際華人統計協會當選理事，同時也是亞洲唯一獲德國洪堡基金會頒發洪堡研究獎的統計專家。

朱教授在多個統計學研究範疇上的貢獻也備受推崇，當中包括降維方法、擬合優度檢驗和經驗似然等－這些都是教普通人如丈八金剛、摸不著頭腦的學術課題。

a world that is becoming ever more complex to understand and analyse. Professor Zhu is more than happy to explain why his work is relevant and beneficial to everyday life.

Take dimension reduction, which in the dictionary is defined as the process of reducing the number of random variables under consideration. In the world of finance, for instance, many stocks in the Hang Seng Index generate huge amounts of information, which statisticians call a high-dimension data problem. This means there are too many variables to be studied simultaneously.

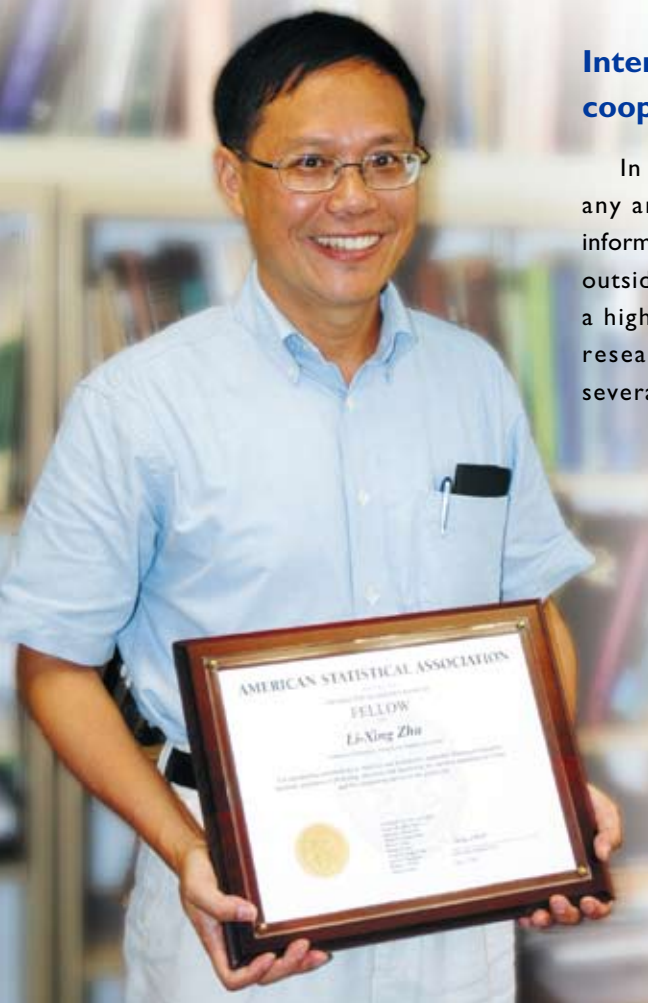
"The results through studying variables one by one may be half-baked," Professor Zhu says. "We have to create new methods to study the variables effectively in order to get the information we need." This is where dimension reduction comes in. The variables, or dimensions, are reduced down to a figure that can be easily disgested to provide the required information.

Another field where dimension reduction is employed is medical science. The human body is highly complicated and collecting data about the human body can be very costly. Dimension reduction is a valuable tool in analysing this data.

Professor Zhu pointed to the use of dimension reduction in analysing the data produced by Hong Kong's Census and Statistics Department each year. The method can also be used in education to analyse students' performance in school to find out which subjects they should study and which university department they should enrol in.

Inter-disciplinary cooperation

In fact the method is valuable in any area where masses of data and information are created – and not just outside the University. HKBU places a high priority on inter-disciplinary research, in which academics from several departments work together,



Professor Zhu Lixing of Mathematics Department was selected as Fellow of the American Statistical Association in 2007
數學系朱力行教授獲選為二零零七年度美國統計協會院士

但事實上，這些研究範疇具有眾多的實用價值。世界發展日新月異，面對難以理解甚至無法分析的社會狀況，這些研究正大派用場。問及他的研究工作如何與日常生活沾上邊時，朱教授便與我們分享統計學的實踐與妙用。

以降維方法為例，字典上的定義是「減少需要考慮的隨機變數的過程」。譬如，在金融市場上，恒生指數眾多成份股都會衍生大量資訊，統計學者稱之為「多維數據問題」，意即同一時間需要研究太多變數。

「逐一研究每項變數，所得到的結果常常是很不完整的。」朱教授說：「我們必須找出有效的方法去同時研究這些變數，才可取得所需資料。」利用降維方法，我們可以把變數(或維度)減至能夠應付的水平，從中擷取所需資料。

另外，醫學研究也經常應用上降維方法。人體構造錯綜複雜，且蒐集人體數據的成本相當高昂，降維方法也是分析這些數據的重要工具。

朱教授又指出，政府統計處的數據有時也可用降維方法分析。此外，在教育方面，降維方法有助分析學生的校內成績，藉以找出他們適合就讀的學系和科目。

跨學科合作

大學十分重視跨學科的研究工作，並鼓勵不同學系的學者攜手合作，集各家所長，從不同角度進行研究。由於降維方法適用於所有牽涉大量數據和資料的範疇，因此大學經常把這統計工具應用於跨學科的研究上。

舉例說，浸大中醫藥學院在進行臨床試驗前，除了要決定安排多少名病人接受特別治療外，還要訂定分析數樣的方法，朱教授也將幫助做這項研究。

另外，朱教授亦與圖書館合作，嘗試找出圖書館使用量與學生的學業成

Professor Zhu conducts statistical research for the School of Chinese Medicine and the Library

朱教授為在中醫藥學院和圖書館進行統計研究

pooling their expertise to tackle research projects from their different perspectives.

Professor Zhu says he has been working with HKBU's highly respected School of Chinese Medicine. When the School wanted to carry out a clinical trial, decisions had to be made on how many patients would be assigned to receive a special treatment and how to analyse the resulting data. Professor Zhu will assist the School in this important work.

He is also collaborating with the HKBU library, which wants to know how important the library is to students' academic performance, as measured by their Grade Point Average (GPA). "Up to now we have used the GPA as one variable, an indicator of academic perfomance. We also used the number of books borrowed from the library as another variable. We wanted to check the relationship between the two."

At first one might predict that students who read a lot of books were, at least, hard workers, and this would reflect in their GPA. But the analysis by Professor Zhu showed that this was not necessarily the case – there was not an absolute correlation between the number of books borrowed and GPA performance.

Be cautious with numbers

Like all statisticians, Professor Zhu is aware of the phrase "Lies, damned lies and statistics." This well-known saying is attributed to 19th Century British Prime Minister Benjamin Disraeli and popularised in the U.S. by author Mark Twain. The statement refers to the persuasive power of numbers, and succinctly describes how even accurate statistics can be used to bolster inaccurate arguments.

Professor Zhu agrees that one must always be careful with statistics. "I always tell people not to believe statistics completely. Why? Because if you use statistics wrongly then statistics are a lie.

"What statistics can do is to provide information, provide a suggestion of probability. We cannot say that this information is 100 percent correct, not absolutely true. If somebody tells you that the information is 100 percent correct then that person is either not a statistician or he might be cheating you."

績和平均積分點(GPA)的關係。朱教授說：「我們希望找出反映學生學業成績的GPA，與借書量這兩項變數之間的關係。」一般人也以為飽覽羣書的學生，總稱得上用功勤奮，GPA自然較高，但朱教授卻發現事實並非如此 — 借書數量與GPA沒存著必然關係。

小心數字

與其他統計學者一樣，朱教授深諳"Lies, damned lies and statistics" 的道理 — 這源自十九世紀的英國首相迪斯累里(Benjamin Disraeli) 的至理明言，由文學泰斗馬克吐溫(Mark Twain)於美國發揚光大，它把謊言和數字相提並論，提醒了人們不應盲目相信數字，因為只要有準確的統計數字支持，即使謊言也顯得可信。

朱教授同意必須小心處理統計數字，他說：「我常叫人不要盡信統計數字。為甚麼？因為只要手法不當，統計數字便無異於謊言。」

「統計學可以提供資料，提出某些事情發生的可能性，但我們無法斷言資料是絕對正確的。假如有人告訴你某項資料百分之百正確，對方縱使不是撒謊，也絕對不會是個統計師。」