Letter from the President

Welcome to this first edition of the ENBIS Magazine. For some time, the European Network for Business and Industrial Statistics (ENBIS) has discussed the prospect of launching a publication to share information among members. In August 2003, a small group was formed to investigate the possibility of starting a magazine that would have a broad appeal to ENBIS membership as well as an outreach towards users and potential users of statistics.

The group, headed by Tony Greenfield, soon returned with an offer from the publisher of Scientific Computing World to publish the ENBIS magazine as part of Scientific Computing World. Considering that Scientific Computing World's editorial content and readership profile agrees very well with the ENBIS vision, the council of ENBIS was happy to accept this offer.

On behalf of ENBIS, I want to thank Scientific Computing World for providing this opportunity to share our information not only among ourselves, but also with SCW readers interested in using contemporary statistical methods in their professional work.

Thanks also to Tony Greenfield and his editorial team for volunteering to do the editorial work.

I am confident that ENBIS members will benefit not only from the ENBIS magazine pages, but also from the reports offered in the main part of the magazine. It is our vision that the wider readership of Scientific Computing World will also find value in the information provided in the ENBIS magazine. My welcome therefore extends to both groups of readers.

Poul Thyregod
President of ENBIS

ENBIS
European Network for Business and Industrial Statistics
www.enbis.org

SEPTEMBER IN COPENHAGEN

The fourth annual meeting of ENBIS will be in Copenhagen,
20 to 22 September 2004.
Keynote addresses will be given by:

- Tim Davis, former Quality Director of the Ford North American Truck Business, who will speak on 'Science, Engineering and Statistics';
- Eva Holm, from Greater Copenhagen Authority, who will speak on 'Monitoring customer satisfaction for bus services in the Greater Copenhagen area'; and
- Jaap van den Heuvel, General Director of the Red Cross Hospital in Beverwijk, who will speak about 'The implementation of Six Sigma in a hospital'.

The Box Medal for outstanding contributions to industrial statistics will be awarded. Many presentations contributed by statistical practitioners will demonstrate statistics in action in a wide range of business and industries.

On 19 September, there will be a pre-conference tutorial workshop on consultancy skills, and on 22 to 24 September there will be a post-conference course on Six Sigma.
Further information can be found on www.enbis.org

ENBIS: a network for discussion

By ***Tony Greenfield***

Industrialists and businessmen across Europe know that international competition is getting tougher, product-development cycles shorter, and manufacturing processes more complex. Their customers expect higher quality in their products.

Statisticians across Europe know that statistical methods have improved business and industrial performance and can continue do so in the future too. But many companies remain ignorant of how they can benefit from the application of statistical techniques. Witness the reply of one company director, when I asked: 'What do you do about uncertainty and variation?' His response: 'They are not allowed.'

The European Network for Business and Industrial Statistics (ENBIS) has been founded to stimulate the application of statistical methods to economic and technical development and to business and industry across the whole of Europe. ENBIS is intended as a forum for the exchange of ideas and to provide a networking mechanism for statistical practitioners. We aim to stimulate the application of statistical methods to enhance economic and technical development and to improve competitiveness of business and industry across the whole of Europe.

Business and industrial statisticians from national societies, such as the Royal Statistical Society in the UK, have been working within companies, or as consultants, to help businesses cut costs and improve quality. Now the availability of the internet and easy travel within Europe has provided the opportunity to create a new society, ENBIS.

The need for networking arose from the realisation that many applied statisticians and statistical practitioners work in environments where they are isolated from interactions with, and stimulation by, likeminded professionals.

ENBIS was created by a small band of enthusiasts and had its official launch in December 2000 at the University of Amsterdam, with more than 80 attending. The provisional executive committee had 14 members drawn from eight European countries. This first meeting had six interest groups: Design of experiments; General statistical modelling; Data mining and data warehousing; Process modelling and control; Reliability and safety; Quality improvement. Since then, two more interest groups have been formed: Statistical consultancy; and Measurement uncertainty.

The European Network for Business and Industrial Statistics was launched in Amsterdam in December 2000.

The first meeting was followed by a three-day course on design of experiments. This is a pattern that has been followed at the annual conferences, the first of which was in Oslo 2001 and was followed by a three-day introductory course on Six Sigma. Courses presented before and after the second and third annual conferences, in Rimini and Barcelona, covered Evaluating the Effectiveness of Advertising, consultancy skills and Six Sigma There will be more at the next conference in Copenhagen 20 to 22 September 2004.

Since that first meeting, solid foundations have been built: membership has grown to nearly 1,000 from more than 30 countries (mostly European but a few from the USA); a constitution has been written; financial management has been created; a secretariat established; and, most visibly, a website has been developed, **www.enbis.org**. It carries reports of the interest groups; information about future meetings and other activities; a report of the first conference; membership details and how to join; and a list of members of the executive committee.

The website has a 'members only' section. This includes a discussion page, enabling members to post messages and reply to others. There is also a network page, which displays all ENBIS members with their affiliations and email addresses. Members can also show further information about themselves, such as their special interests and areas of expertise. Each interest group has its own discussion page to exchange ideas; all members can join in.

Membership of ENBIS, and hence of its interest groups, is open to engineers, scientists and managers working in business and industry who use statistical methods in their work. The discussions and information exchanges through the interest group pages will develop into valuable resources for these 'statistical practitioners' as well as for the professional statisticians who are also eligible for membership.

Businesses and industries of Europe do need ENBIS. Some of them know this; many do not. Our task is to reach out and show how statistics can help them succeed in their own enterprises.

Tony Greenfield is chairman and chief consultant of Greenfield Research Ltd, visiting professor to University of Newcastle upon Tyne, and a past president of ENBIS.

Banking risk? Rely on Bayes

***Paulo Giudici** suggests that a Bayesian data mining model could allow banks to manage their operational risks*

Operational risks in banking have increased in recent years. These risks spring from globalisation of the financial markets, growth of IT, and the spread of complicated financial products. The new Basel Capital Accord requires banks to put aside a minimum capital sum to insure against credit risk, market risk, and operational risk. Of this sum, 12 per cent is specifically assigned to cover operational risks. The Basel agreement is still being revised, but is expected to become compulsory for all banks from 1 January 2007. (A press release with more details can be found at www.bis.org/press/p040115.htm).

Operational risk is defined as 'the risk of losses resulting from inadequate or failed internal processes, people and systems or from external events'. It includes legal risks, but not strategic or reputational risks. In the new version of the Basel Agreement, operational risks have been split into a series of standardised business units, called 'business lines', and into groups according to the nature of the underlying risk, called 'event types'.

The Basel Agreement allows larger banks to follow Advanced Measurement Approaches (AMA). In such approaches, the amount of capital required by the regulations will equal the risk measure generated by the bank's internal system for measuring operational risk, using quantitative and qualitative criteria set by the Basel committee.

As a consequence of the Basel Agreement, banks must predict operational risks accurately, as well as credit and market risks. The data-mining laboratory of the University of Pavia (www.dataminginglab.it) has recently proposed an approach, based on Bayesian network models, that seems to be effective in predicting and managing operational risks. This approach has been developed and validated according to the general data mining principles set out in Giudici, *Applied Data Mining* (Wiley, 2003).

One of the main problems in measuring operational risks is the scarcity of data on internal losses. Our proposed methodology solves this critical point because it allows a coherent integration, via Bayes's theorem, of different sources of information, such as internal and external data, and opinions of 'experts' (process owners) about the frequency and the severity of each loss event. Furthermore, the proposed model corrects the loss distributions by considering possible causal factors ('key risk indicators') as well as possible relationships between different nodes, representing combinations of the business line/event type/bank/process.

We successfully tested the methodology in a pilot project developed for the Italian banking group, Monte dei Paschi di Siena (MPS). The project is now being developed and consolidated, both in methodological and computational terms.

The Bayes approach

One of the main problems about operational risk is the absence of an appropriate historical database. This makes it difficult to estimate operational risk through classical statistical inference. A 'self-assessment' approach may alleviate the lack of historical data, in that expert opinion can provide a rough estimate of predicted operational losses. These expert opinions must, however, be integrated coherently with actual loss data.

We have proposed, using Bayes's theorem, to combine qualitative data (opinion) with quantitative data, internal and external. We chose a Bayesian network model to describe explicitly the correlation between losses, so as to measure the impact on them of various causal factors, such as the efficiency of control systems, or key risk indicators like computer systems load, and personnel at work.

Our proposed Bayesian network model has three main types of variable nodes.

- Losses: for each combination of business line/event type/company/process levels, according to Basel requirements, we have loss data at the process level. As each company within the banking group can have different size or business volume, for each combination we have 26 possible states.
- Effectiveness of internal controls (ic) on each business line /event type/company/process, with the following states: N/A = no opinion expressed; 1 = no existing; 2 = no effective; 3 = need adjusting; 4 = effective;
- Effectiveness of external controls (ec) about each business line / event type / company/process with the following states: N/A = no opinion expressed; 1 = no existing; 2 = no effective; 3 = need adjusting; 4 = effective.

For each loss distribution in the Bayesian network, we can calculate any percentiles of interest, such as the 99.99 percentile required by the Basel committee (Value at Risk). If we estimate the number of future losses in a given period, such as one year, we will be able to calculate the capital to be put aside by the bank: the 'operational Value at Risk' which is the VaR per event multiplied by this estimate of frequency.

In this way, we obtain one operational risk VaR for each node. By the nature of the Bayesian network, we can also easily determine aggregated business line/event type operational risk VaR, summing the VaR for every society and process for each business line/event type combination.

The resulting overall VaR OP represents the capital charge required to satisfy the first pillar of the new Basel Capital Accord.

In this case, it is simply the sum of the various VaR, as correlations between losses are already taken into account by the model. This is in contrast with what happens, for instance in the mainstream actuarial approach, (see Cruz, *Modelling, measuring and hedging operational risk*, Wiley, 2002), which assumes independence between the different nodes.

Paolo Giudici is at the University of Pavia.
This research was supported by the pro-ENBIS network under contract GTC-2001-43031.

Better insulin injectors from statistical experiments

A statistical analysis showed **Peter Thyregod** *and his colleagues how to improve the quality of insulin injectors for diabetics*

Figure 1 The seven moulded parts that go to make up the injector.

Over the past few years, demand for injection-moulded products has increased. In the past, the moulding process concentrated on high-volume products with low to moderate quality requirements. In recent years, the market has expanded to include precision-moulded items, such as medical devices or connectors for the auto industry, which have a higher quality requirement.

Designs for plastic parts are becoming more complex, while economic constraints have led to the use of multi-cavity moulds that produce several identical parts in each machine cycle. The number of dimensions critical to functionality of a moulded product is increasing, not only as a result of part geometries, but also with the number of cavities in the mould.

A particular medical device, a pen injector used by diabetics for injection of insulin, is assembled from seven parts, all moulded in plastic (shown in Figure 1). The functionality of the device depends critically on how well the parts fit together. Consequently the quality requirements for each part are very strict: they allow for little variation in the critical dimensions. A medical device that does not meet the requirements should not reach the consumer, and so both moulded parts and final assembled products need to be inspected. If a part fails the inspection, costly corrective actions, such as screening to delimit unsatisfactory production, have to be taken. The benefit of our work to the manufacturer was a considerable reduction in process variation, and hence a substantial cut in such screening.

For this experiment, we chose the one part that most often failed inspection. The critical dimensions of this part were length, and inner diameter at each end. Random variation in the production process leads to some probability that some parts will not fit together properly. Even small deviations from target will lead to parts that do not meet specifications. A good experiment would identify the best settings of process control variables to minimise the variabilities of critical dimensions, and hence minimise the probability of a faulty assembled pen.

The manufacturer believed that the mould itself influenced the magnitude of the differences between cavities, and that the settings during the injection-moulding process itself had little effect. On the other hand, the operator (the moulder) did use machine variables to control dimensions such as part lengths. The operators knew that some dimensions were influenced by machine settings. We undertook a study to check this assumption on the manufacturer's part. Eventually, we were able to show that several of the settings did indeed influence the quality of the product. We could then advise the manufacturer on how to control quality through the settings.

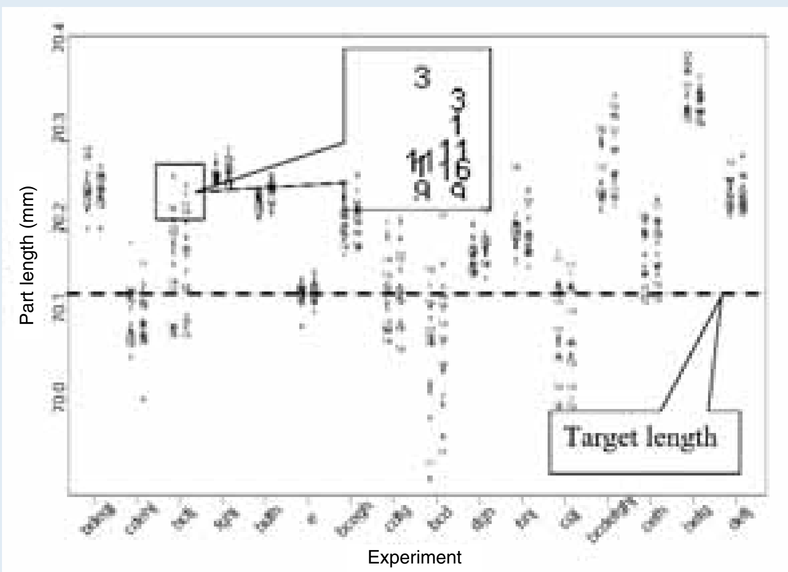

Figure 2 The measured length of the moulded parts in each of the 16 experiments. The numbers 1 to 16 are the cavity numbers. The horizontal line denotes the target length.

The experiment

We set up an experiment to assess the influence of different machine settings on part quality. The process engineers and the statistician jointly made decisions about the experiment. They chose a hollow cylindrical part, produced by a 16-cavity mould for the experiment. The main quality parameter was part length. They selected nine machine variables, based on experience and injection moulding theory. The nine variables are shown in the Table opposite. The levels were chosen such that they were clearly outside ordinary operating conditions but not further than still allowing for a useable product.

To assess which variables and combinations influence mean length and cavity-to-cavity variation in part length, a fractional factorial experiment was performed corresponding to 16 selected combinations of the nine machine variables.

From each of the 16 different combinations of parameters, two shots were collected and measured. Each shot has 16 parts, one from each cavity. The results are displayed in experimental order in Figure 2. The numbers 1 to 16 are the cavity numbers. For each experiment the letters on the x-axis denote the standard factorial experimental points: for example, bhj indicates that variables B, H and J were each at their high levels, while all other variables were at their low levels. Each cavity number appears twice at each experimental point.

The findings

Figure 2 reveals that both mean part length and cavity-to-cavity variation are greatly affected by the different settings. An analysis of variance further confirmed that changes in settings in four of the variables (holding pressure, injection time, holding pressure time and back pressure) lead to significant changes in mean length. This result is in good agreement with process experience, and two of the machine variables are usually used by the moulder to control part lengths.

Furthermore, another analysis of variance revealed that two of the variables (injection time and cooling time) had significant influence on the cavity-to-cavity variation in part length. This was not anticipated by the process engineers.

As injection time is among the variables most often used for controlling the mean length of the part, this result has wide implications for the control of the process. In the past, when attempting to adjust the mean length by adjusting injection time, an unwanted side effect was an increased cavity-to-cavity variation. Further analysis of the data suggested a setting where the cavity-to-cavity variation is kept at a minimum, so that the process is more robust to variations in raw material and environmental conditions.

Henrik Madsen and Henrik Spliid are with the Technical University of Denmark while Peter Thyregod is with Novozymes and Henrik

Variable		Low setting	High setting
A	Mould temperature	Cooling	No cooling
B	Holding pressure	53 bar	63 bar
C	Injection time	0.3 sec	1.6 sec
D	Fill/post-fill switch	12.5 mm	16.5 mm
E	Barrel temperature profile	(205, 215, 205, 185) ºC	(235, 245, 235, 215) ºC
F	Holding pressure time	2.5 sec	4.5 sec
G	Back pressure	5 bar	30 bar
H	Cooling time	2 sec	6 sec
J	Travel	47. 5 mm	51.5 mm