LETTER FROM THE PRESIDENT

Facing the future with confidence

www.enbis.org

By Shirley Coleman

I have greatly enjoyed being president of ENBIS for the past year and getting to know many of the members. I was previously familiar only with the quality improvement special interest group, which I chair, but now I know about all 10 groups. Lots of action goes on behind the scenes to develop the network. We are a young organisation and are learning and changing every year. The executive committee – three vice-presidents, the current and the next president – met several times during the year, both face-to-face and electronically. We want to find out what members think of ENBIS, and to know if there are members in key organisations in each country. We realise that the national representatives have a vital role to play in promoting and enlivening ENBIS, and so we started by making contact with them. We would like to survey the membership using questions in their native language. We have asked national representatives to list five main organisations in their country, so that we can see which of these organisations have ENBIS members and where to promote ENBIS. The new executive committee has already met to plan its activities and, as the past president, I shall also take part.

A record number of people attended the fifth ENBIS conference in Newcastle. It kicked off with an excellent keynote presentation, by Doug Montgomery of Arizona State University, about the role of statistics in the transformation of science, business, and industry. We received a record number of abstracts, and there were three or four parallel sessions each day. There were many opportunities for networking during the conference, particularly in the Magpie Suite at St James's Park, overlooking the football ground.

This year, there were nine workshops before, during, and after the conference, arising from the special interest groups. These were well attended and had very good reviews. One idea from the pro-ENBIS project, which was successfully completed earlier this year, was to set up an ENBIS academy. These nine workshops were the ENBIS academy in action: they dealt with data mining; operational risk management; simulation in clinical trials; statistics in innovation and the design process; research methods; reliability; advanced methods; the wild river designed experiment; and statistical consulting skills.

The Magpie Suite at Newcastle's St James Park was a good venue for networking.

ENBIS members always seem to be determined to enjoy themselves and to get the best from every event, and there were many positive comments about the conference. One of the highlights for me was the conference dinner. We invited Robin Plackett as our guest of honour. Robin was the foundation professor of statistics at Newcastle University, and worked at the University for many years before he retired in 1983. Robin is famous for many statistical achievements, including the Plackett-Berman designs that are widely used in industrial experiments. He published the designs in 1946 as an interesting piece of theory. Recently, with the increasing interest in designed experiments in industry, the designs have become widely used in product development and quality improvement. As it happened, it was Robin's 85th birthday and it was good to celebrate with him. Robin worked with Peter Armitage, of Oxford University, in operations-research during World War II, and they have enjoyed a long friendship. Peter, who works mainly in medical statistics, gave an excellent after-dinner speech remembering the good times that he and Robin shared.

The George Box medallist was Sir David Cox. Before his academic career, he worked with WIRA (Wool Industry Research Association). He told the conference about the complications of wool fibre, how it undulates as a consequence of the variation in growth, and can be smoothed in only one direction. It is inspirational for young statisticians to hear tales from the great shapers of statistics and to meet the people whose names are attached to many of the techniques in everyday use.

ENBIS conferences are an effective way for everyone to meet and share ideas. A mid-term ENBIS meeting is planned for March 2006, in advance of the sixth ENBIS conference, which will be held in Wroclaw, Poland from 18 to 20 September 2006. The seventh conference will be in Dortmund, Germany. The network is in good shape, and ready to flourish this year and beyond. The website, www.enbis.org, has details of the conferences and other meetings. I look forward to seeing you at these events.

This is Shirley Coleman's last President's Letter. She has now handed over the presidency to Fabrizio Ruggeri, of the Istituto di Matematica Applicata e Tecnologie Informatiche, Milan, Pavia, and Genoa.

Estimating people's values

Roselinde Kessels, Peter Goos, and Martina Vandebroek are designing experiments to analyse how consumers value new products

Attribute		Level 1	Level 2	Level 3
1	Range (km)	22	32	35
2	Charging time(hours)	2	3	4
3	Maximum speed (km/h)	25	32	60
4	Price (?)	800	2500	3000

Range: up to 22 km	**Range:** up to 32 km	**Range:** up to 35 km
Charging time: 2 hours	**Charging time:** 3 hours	**Charging time:** 4 hours
Maximum speed: 25 km/h	**Maximum speed:** 32 km/h	**Maximum speed:** 60 km/h
Price: 800?	**Price:** 2500?	**Price:** 3000?
Preference rating: 1-2-3-4-5-6-7-8-9-10	**Preference rating:** 1-2-3-4-5-6-7-8-9-10	**Preference rating:** 1-2-3-4-5-6-7-8-9-10

Figure 1: Composition of profiles of an electric bike for a rating-based conjoint experiment.

Suppose that a motorbike manufacturer wants to increase its market share by adding several electric bikes to its product range. Because the company is interested in making a profit, it wants to produce electric bike configurations that appeal to a large number of people. One way to learn about consumers' preferences would be to ask respondents to state how much they value different levels of some of a motorbike's attributes. This is difficult for the respondent.

Conjoint analysis, on the other hand, attempts to simplify the respondents' task by letting them evaluate several alternative configurations. For the motorbike example, (see Figure 1) these alternatives are described in terms of combinations of attributes. In choice-based conjoint studies, the evaluation solely consists of choosing the most appealing alternative. In rating-based conjoint studies, the respondents have to provide a rating for each alternative presented to them. The respondents' answers are then converted into importance values, utilities, or part-worths for the attribute using a statistical model (a non-linear model for a conjoint-choice experiment, and a linear one for a rating-based study).

Now assume that the motorbike distributor is willing to construct several prototype motorbikes, say 27 of them, and to conduct a rating-based conjoint experiment in which some potential consumers can try out several motorbikes over a couple of weeks. The manufacturer then immediately faces the question of which motorbikes to construct. Another issue is how many respondents or test persons are needed. Finally, the manufacturer needs to decide how to assign the 27 prototype motorbikes to these respondents. As shown in Figure 1, each motorbike in the study is defined by four attributes (range, charging time, maximum speed and price), and each attribute acts at three levels. As a result, 3 x 3 x 3 x 3 = 81 different motorbikes can be described using the four attributes. The figure shows three of these motorbikes.

The motorbike experiment demonstrates that in a rating-based conjoint experiment several alternatives have to be presented to a limited number of respondents. The alternatives presented to the respondents are not always tested in conjoint studies: for reasons of cost, the alternatives are often merely verbal descriptions, sometimes combined with a graphical representation. The aim is to elicit information on the utility the respondents derive from the levels of the products' or services' attributes. These utilities contain information about consumers' trade-offs, as well as predictions about their future behaviour.

The quality of these inferences depends on the alternatives and the number of test persons used in the study. The assignment of the alternatives to the subjects also plays a key role. This is where experimental design can play a crucial role.

Experimental design techniques can answer the questions:

- What are the best alternatives to use in a conjoint study? and
- How should these alternatives be presented to the respondents?

A statistically efficient experimental design is needed if conclusions are to be reliable and the part-worths estimated precisely. The literature on design of conjoint experiments is, however, silent about how to select sets of alternative products or services to be evaluated by the respondents. In the motorbike conjoint study, the literature provides a tailor-made answer neither on how to select 27 prototypes out of the set of 81 possible ones, nor on the ideal number of test persons, nor on the partition of the 27 selected alternatives over those respondents. It is, for example, unclear whether it is better to submit nine alternatives to three respondents or three alternatives to nine respondents, or even to have 27 respondents evaluate just one motorbike each.

The answers to such questions can be found by applying the techniques of optimal design of experiments. As the type of experiment considered here involves tests or observations in groups, its design is similar to that of 'blocked experiments'. Blocked experiments are used in industry and agriculture, when the observations are obtained in groups (called 'blocks' in the literature) and when the observations within each group are more similar than observations from different groups. In industry, such situations occur when more than one batch of material is required, or when the experiment takes up more than one day. In conjoint studies, it is also quite likely that test results from one respondent will be more similar than results from different persons. Respondents are heterogeneous because they are all different with respect to age, experience with the product or service under study, physical characteristics, cognitive abilities, and preferences. In statistical terms, the observations from one respondent are positively correlated.

Statistical models corresponding to blocked experiments explicitly take into account the correlation pattern of the data because this allows a more powerful inference. To find the best possible setup or design of the experiment, it is also good to take the correlation structure into account. We have done so in designing the motorbike conjoint study. To find the best possible design, we used an algorithmic approach to compute the so-called D-optimal design. In theory, these optimal designs depend on how similar the responses are from a single respondent (or how different responses are from different subjects). Fortunately, the optimal designs were not sensitive to the resulting degree of correlation.

The algorithmic construction of the optimal designs for the motorbike experiment involving 27 motorbikes showed that, from a statistical point of view, it is most efficient to have three motorbikes rated by each of nine respondents. The best possible design for the motorbike example is displayed in Table 1. If the number of prototypes a researcher is willing to construct is not a multiple of three, all but one respondent have to rate three bikes.

A problem with the algorithmic approach is that it requires a long time to find the optimal conjoint designs when a large number of prototypes is available. A possible solution might be to construct a conjoint design for a simpler situation with a smaller, but not too small, number of prototypes and to use that design a couple of times. This approach yields much better results than approaches in which several prototypes are selected while ignoring the correlation structure and assigning the selected prototypes at random to the respondents. Figure 2 shows that the latter procedure leads to substantial losses in information for the electric bike example. Many more prototypes are required when the correlation structure is ignored and the prototypes are randomly assigned, as compared to an experiment that is optimally designed. The three coloured lines in the graphs correspond to situations where there is not much correlation between observations from a respondent (green line), a moderate correlation (orange line) and a large correlation (blue line). For a moderate degree of correlation, an optimal design with only 32 prototypes provides as much information about the part-worths as a conjoint study with 40 prototypes that was not optimally designed.

HOW TO ESTIMATE VALUE

Conjoint analysis encompasses a broad range of techniques for estimating the value people attach to the attributes or features of products and services. Roughly speaking, a conjoint analysis comprises the following seven steps:

1) Determine the attributes of the product or service that might be important to consumers, as well as the levels of these attributes.

2) Decide on the data collection methodology to reach respondents, such as mail, telephone or internet.

3) Select the conjoint technique that best fits the research problem: choice-based conjoint studies require the respondents to indicate their choice for one out of a set of products or services: rating-based conjoint studies require the respondents to rate a series of products or services.

4) Select the products or services (these are nothing but combinations of attribute levels) presented to the respondents. This is usually called the 'experimental design problem'.

5) Collect the data.

6) Estimate the utilities attached to the attributes' levels. These utilities are often referred to as part-worths and are parameters in a statistical model.

7) Use the estimated statistical model for prediction purposes.

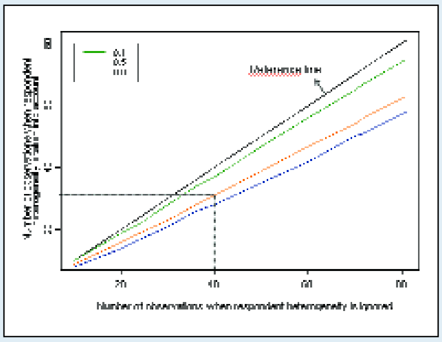


Figure 2: **Iso-information lines for the prototype conjoint study. On the horizontal-axis: Number of prototypes needed when the conjoint study is not designed using optimal design methodology. On the vertical axis: Number of prototypes needed when optimal design methodology is used. The green, orange and blue lines correspond to small, moderate and large degrees of correlation, respectively. The black line is the bisector and serves as reference line.**

Roselinde Kessels is a doctoral student at the Katholieke Universiteit Leuven, Belgium (roselinde.kessels@econ.kuleuven.be). Peter Goos is professor of statistics at the Universiteit Antwerpen (peter.goos@ua.ac.be), and Martina Vandebroek is professor of statistics at the Katholieke Universiteit Leuven (martina.vandebroek@econ.kuleuven.be). Readers who would like more background regarding the technical aspects of the problem are invited to contact one of the authors.

	Attributes					Attributes					Attributes			
Resp	**1**	**2**	**3**	**4**	**Resp**	**1**	**2**	**3**	**4**	**Resp**	**1**	**2**	**3**	**4**
1	2	1	1	1	4	1	1	2	1	7	3	1	3	1
1	1	2	2	2	4	3	2	3	2	7	2	2	1	2
1	3	3	3	3	4	2	3	1	3	7	1	3	2	3
2	1	2	1	1	5	3	2	2	1	8	2	2	3	1
2	3	3	2	2	5	2	3	3	2	8	1	3	1	2
2	2	1	3	3	5	1	1	1	3	8	3	1	2	3
3	3	3	1	1	6	2	3	2	1	9	1	3	3	1
3	2	1	2	2	6	1	1	3	2	9	3	1	1	2
3	1	2	3	3	6	3	2	1	3	9	2	2	2	3

Table 1: **Optimal conjoint design in which nine respondents each rate three prototypes.**

Schools of statistics

John Logsdon invites readers to throw some virtual dice to learn about different schools of statistics.

Statisticians can be quite a dynamic bunch – and amusing dinner guests. So it should not surprise anyone to learn of internecine statistical arguments, where much blood is spilled on the conference or seminar carpet.

There are three schools of statistics: classical, frequentist and Bayesian. How do these relate to each other? This article is the first of two, trying to lift the curtain on these obscure terms.

Simply put: in classical statistics there is no data at all; a frequentist adds only the data; a Bayesian adds some belief or prejudice. To illustrate, I have constructed an experiment – it can be found at www.enbis.org/SCW1.R. It requires the program R (see box). The rest of this article will go through the experiment, as an interactive guide to basic statistics.

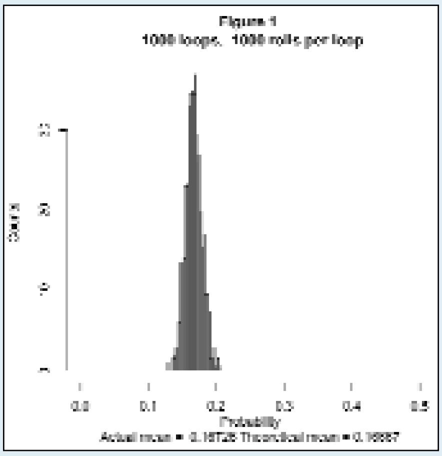
Windows users should click the icon and select 'Source R code' from the File drop-down. In Unix, just type R at the prompt then type: source(filename), where filename is the name of the file containing the downloaded code. Windows users can browse for the file.

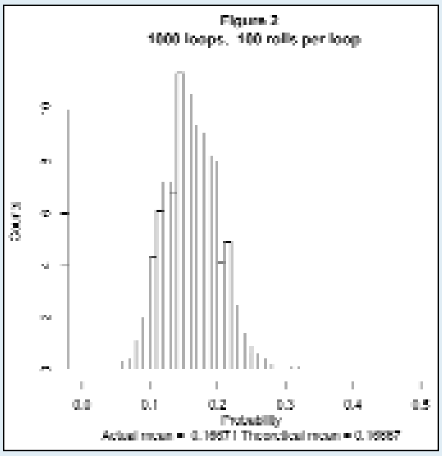
There are three functions in this file: class.calc, freq.calc and bayes.calc. – each based on throwing dice. Here, I will deal with the first two; the last will be discussed in the next issue.

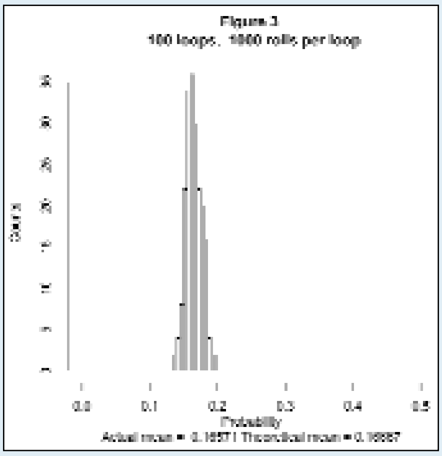
If the function is run with no parameters, it will explain the meaning of the parameters. The arguments can be specified by position, parameter, or a mixture.

class.calc() returns the classical result and gives the same result each time – there is no data to confuse it. Tell it that there are six sides to your dice – by typing class.calc(n=6) – and it will tell you that the probability of any particular side coming up is 0.1666667. It always returns the same answer for the same number of faces. This may seem pretty uninteresting, but classical statistics provides the foundations for statistical modelling.

freq.calc() simulates some numbers and plots a histogram of the probability that a number is rolled. A full command might be: freq.calc (n=1000,m=100,face=1,up=6,xlow=0,xupp=1)







The R program can be downloaded from www.r-project.org. It is the open-source software implementation of the S language, also used by S-Plus. It has a vast library of commands and specialist add-on functions.

There are six parameters but it is the first two that are interesting: m is the number of throws necessary to estimate the probability of a given face value; n is the number of times this is tried, so you can work out a distribution of the probability. You can set up to the number of faces (default is 6) and face to any number from 1 to the number of faces (default is 1). You can also specify lower and upper limits to the histogram.

You can vary the number of rolls of the die within the loop, which estimates the individual means, and the number of loops, which generates the distribution. For example, you can play with the number of rolls of the die you need to estimate the probability that a one comes up, and then with the number of such exercises.

If m is small there will be a lot of variation, since the probability is just the number of ones divided by the number of rolls. It is even possible that there are no ones at all. You need to ensure that there are enough results for the histogram function to plot the data or you will get an error. A typical result is given in Figure 1.

Try with only a small number of rolls and it will give a very large spread, however many loops you try. With a large number of rolls, the answers will be more consistent. It makes no difference which face you select – the answers will be different only because you have selected a different set of probabilities.

Figure 2 shows what happens when there are a lot of rolls but few samples taken from this distribution. Figure 3 shows what happens when there are a lot of samples, each of a few rolls. Notice how much wider the distributions are when there are few rolls. To compare one with another, you can set the lower and upper limit: freq.calc(n=1000,m=10,xlow=0,xupp=1)

Just playing with this program is instructive. It is surprising how much spread there is with a large number of both m and n. This is the underlying randomness we all have to live with.

But what do these values mean? There are many answers to this, but here we estimate the value of the parameter by maximising the probability that the data arose from the model.

I will be writing more about likelihood in the next issue and introducing Bayes. Meanwhile, keep the champagne on ice and, as homework, you can play with bayes.calc().