Modelling buyer behaviour/1

The relevance of buyer behaviour models

By Chuck Chakrapani

Implicit models are universal

The area of marketing models is said to be controversial. Yet practically all research and marketing processes involve some form of modelling. Even marketers and researchers who may not consider themselves 'modellers', are constantly involved in some form of implicit modelling. For instance, when a researcher collects data on overall satisfaction and cross-tabulates it with product attributes, s/he is working with an implied model which states that overall satisfaction is a function of product attributes. In mathematical terms this can be expressed as

$$OS = f(x1, x2, x3 \ldots)$$

where

OS = Overall Satisfaction

x1, x2, x3 ... = different product attributes

Anyone who is involved in building formal models is probably trying to assess the extent to which each variable effects overall satisfaction, the nature of the relationships among different product attributes and how they interact with each other and collectively influence overall satisfaction.

The nature of controversy

Given that a formal model builder is simply extending and clarifying the researcher's implied hypotheses, it is surprising that there is any controversy at all about models. After all, the mathematics that is used in the formal modelling process is the result of forcing the researcher to state his or her hypothesis more clearly.

Yet there is a problem. When a researcher uses an implied model, s/he knows what s/he is looking for. In mathematical modelling two additional factors are often present:

1. Since many researchers do not specialize in formal modelling processes, once the initial hypothesis is specified, the researcher/marketer has hardly any control over how the models are arrived at or what alternatives are available to the researcher. If the modelling is handled by an incompetent modeller, it is not easy for the researcher to be aware of the problem.
2. In some cases, the modeller does not even explain the basic weaknesses of the models, as often happens when a marketer buys a commercially marketed proprietary model.

In either case, what the researcher/marketer is buying is a 'black box' content which is either difficult to understand or not revealed.

It is also not uncommon to find people who invent a model (especially for commercial purposes) and then claim mystical properties for these models. Such claims are coupled with claims of predictive accuracy that are unheard of in any other area of human endeavour.

The discomfort that arises out of not knowing is exacerbated when the results of the modelling process lack face validity. Models that are conceptually and methodologically complex are greeted with scepticism.

The relevance of models

It appears that the nature of most objections relate to the 'black box' approach to modelling and to unnecessary complexity. Yet it is obvious that the modelling process is inherent in data

analysis even if one consciously rejects the idea of modelling. If modelling is inherent in the way we interpret data, we can strengthen our understanding by
• clarifying and extending the models implied in our interpretation; and
• make them explicit and formal
As long as the model is simply clarifying, formalizing and extending the thinking of the researcher, it is hard to see how modelling can do any harm. Quite the contrary. By forcing the researcher to clarify his/her assumptions, the modelling process challenges the researcher to be very specific and explicit about those assumptions.

What is a model?

Models are essentially hypotheses of how things work. Most models study certain aspects of reality (purchasing a shampoo and consumer demographics) and the relationships among the variables chosen to represent that reality (how demographics affect shampoo purchase). Formal models tend to express the relationship in an explicit way using mathematical notation. Inferences are then drawn by identifying the relationship among the variables. Most formal models use quantitative data obtained empirically using mathematical manipulation to draw inferences.

Some areas in which modelling is useful

Models have been particularly valuable in the following areas:
• How consumers act (including how they make their purchase decisions, how they read a newspaper and how some of their behaviour is related to their other behaviour)
• What factors are influential in changing their behaviour
• How to assess the future market structure
• What is the likely impact of alternative marketing strategies
• Identifying the gaps in our knowledge
• Improving the gaps between management and researcher/marketers

Model dimensions

Since the word ‘modelling’ applies to any kind of abstraction, it becomes necessary to understand what type of model is being described. Some common types of models are:

Level of generality

A model is *macro* if it is global. For instance, a model that attempts to describe the totality of buyer behaviour as influenced by a variety of characteristics is a macro model. A model is *micro* if it describes a specific aspect of a system. For example, as a result of a survey, if we arrive at a conclusion such that in 1992, there was a predictable relationship between discretionary income and buying a certain make of automobile, what we have is a micro model.

Level of quantification

When a model is described verbally to explain an observed result we have a *qualitative* model. Such descriptions are usually based on assumptions and theories which may or may not be supported by past research. A statistical model, on the other hand, specifies explicitly and precisely what the assumed and proven relationships are. Both inputs and outputs to the model are quantitative. It should be noted that just because a model uses quantitative or statistical techniques it does not follow that that model is more valid (or even valid at all) compared to another model which does not use quantitative procedures. Good models generally combine both statistical and *qualitative* aspects.

Type of relationships studied

Sometimes a model aims to understand the current structure of the market. For example, segmenting the market helps us to understand the structure of the market. Such models are *structural* models. Sometimes we are interested in predicting buyer behaviour. If we know the consumer characteristics in a trade area, can we predict the success of a new retail establishment? Such models are *predictive* models.

In this series, we will explore some models that are based on research and are relevant to buyer behaviour. For illustrative purposes I will confine my discussions to quantitative models of buyer behaviour. The purpose of the discussion is to explore the model building process and understand its advantages and limitations.

Latent and manifest variables

Consider a market research survey in which we obtain data from a consumer panel. The panel reports on their purchase of product A over a 4-week period. What does this data tell us? Surely the actual data itself is subject to a number of vagaries such as the time period during which the study was carried out, the timespan (4-weeks) chosen for the study, the sample size, etc.

A given consumer's reaction could have been influenced by a myriad of variables when the study was carried out. Yet the analysis, to be meaningful from a modelling point of view, should identify some underlying relationship. We need to extract some basic relationship that will strengthen our understanding of buying behaviour and help us predict future behaviour. When buying behaviour turns out to be different, the model should be able to identify the reason for this result.

The relationship we seek tends to be latent. For example, if our interest is in finding out the rate at which a consumer will buy a product given his/her demographic characteristics, we immediately note that rate is not an observable variable. All that the consumers told us was the *frequency* with which they bought a given product. For example the rate of purchase could be .25 times per week. This may or may not mean that the respondent in the panel made exactly one purchase during the survey period. While rates are abstracted and can be generalized, frequencies are observed and not necessarily generalizable. Frequency is a manifest (observed) variable while rate is a latent variable. The latent variable 'rate' expresses itself through the observed variable 'frequency'.

Modelling is useful in identifying latent variables

consumer behaviour is characterized by such latent variables as: the probability of a person voting, a person's probability of arriving late to a meeting, the probability of forgetting to do something, the probability of a person reading a newspaper every day given he/she has post-graduate education - all these and many more vary from individual to individual. yet they are in a way predictable. When we make a statement such as 'Joe misses meetings half the time', we are, however crudely, expressing our estimation of the latent variable 'rate'.

Let us assume that each person has a rate, based on which s/he expresses a particular behaviour (such as buying clothes).

From frequencies to rates

How exactly do we go from frequency of purchases to rate of buying? After all, we already noted that just because a consumer bought two times in a four week period we cannot assume that the rate of purchase per week is .5. A consumer who has a rate of purchase of .25 per week does not have to buy exactly once during our survey period. Then how do we arrive at a 'rate' given the 'frequency' data?

Stochastic models

The class of models that enables us to model time-varying processes is called 'stochastic'. To model the stochastic process that is of interest to us now, we start with the following assumptions.

- Every consumer has an underlying 'rate' over a given span of time.
- Different consumers have different rates.
- The distribution of these rates (when we consider a group of consumers) will follow a definable mathematical function.
- For any given rate, there is a frequency distribution. For example, even though the rate of purchase is .25 per week, during a 4-week period a given consumer can buy 0, 1, 2, ...times. (this process is known as 'stochastic'; it refers to the random error part of the model.).
- for any given rate, the frequency of purchase by a group of consumers can also be closely approximated by some mathematical function.

Combining the two distributions

Now we have two distributions. One, a distribution of rates at which different consumers buy the product and two, a distribution of frequencies corresponding to each rate.

To arrive at a Stochastic model, we identify and combine these two distributions. The model we thus arrive at will belong to a logical group of mathematical curves. The exact curve finally arrived at would depend on the values of the parameters that will be decided by the nature of our data. In fact, the model building process is closely tied to the identification of the parameter values.

Reality check

A model is generally arrived at on the basis of manifest variables. (To generate 'rates', we need 'frequencies'). We expect the model to predict the future frequency of consumer buying. Therefore, we reverse the process. Instead of going from a manifest (frequency) to a latent (rate) variable, we go from rate to predicting frequency. We expect the frequency predicted by the model to match the survey data.

Is the model valid?

How do we know if a model is valid? If we build a model and predict some aspect of buyer behaviour, does that necessarily mean that the model is valid? The answer to this question will depend on several considerations.

If our only interest is predicting a behaviour, a model is valid to the extent it accurately and consistently predicts the behaviour under consideration. For many quantitative buyer behaviour models, this is of primary concern.

If our interest is in not only predicting a behaviour but providing an explanation for its occurrence, the validity of the model depends on its theoretical soundness (preferably supported by empirical evidence).

Some modellers hold that it is not necessary to explain the phenomenon under consideration as long as we can predict a given behaviour accurately. While this assertion has some merit, problems arise when the model breaks down at some point in time. Ideally, a model should provide some explanation as to *why* it works.

Dr. Chuck Chakrapani, is President of Standard Research Systems and author of Marketing Research: Methods and Canadian Practice (with K. Deal).