TOWARDS THE GRAND UNIFICATION : THE PERSONAL PROBABILITIES ROUTE

Gilles Santini, IMS France

1- THE PROBLEM

The major achievement of theoretical physics has been in the second part of the XXth century to find a common framework to unify the four fundamental forces of the universe. Such achievement, known as the Grand Unification, has required tremendous mathematical work and brave conceptual innovation. It has only been completed recently and some major aspects still need clarification and validation.

Nevertheless, it will be considered in the future as one of the major scientific breakthrough in a century where the balance between good and evil is still undecided.

The reason why I got interested in the Grand Unification process, despite my usual inclination for some of the unchartered lands of mathematics, is that it may give us some leverage to achieve a better understanding of the concept of audience.
In order to get to this point, I need to give some exploratory details first.

As most of you should know there are four fundamental forces in the universe :

- The electromagnetism
- The gravitation
- The weak interaction
- The nuclear force

Electromagnetism is responsible for electronic and magnetic phenomenons. The gravitation governs planets and galaxies attractions. The weak and the nuclear force rule subatomic interactions.
When searching for a unified theory that can account for all of these four forces at the same time, the physician is faced to a dilemma.

On one hand, everything that happens at subatomic level is well predicted by Quantum theory, a theory which is not very intuitive but nevertheless provides a fair and validated description of the subatomic world. On the other hand, gravitation is well understood and accounted for by Einstein Generalised Relativity Theory.

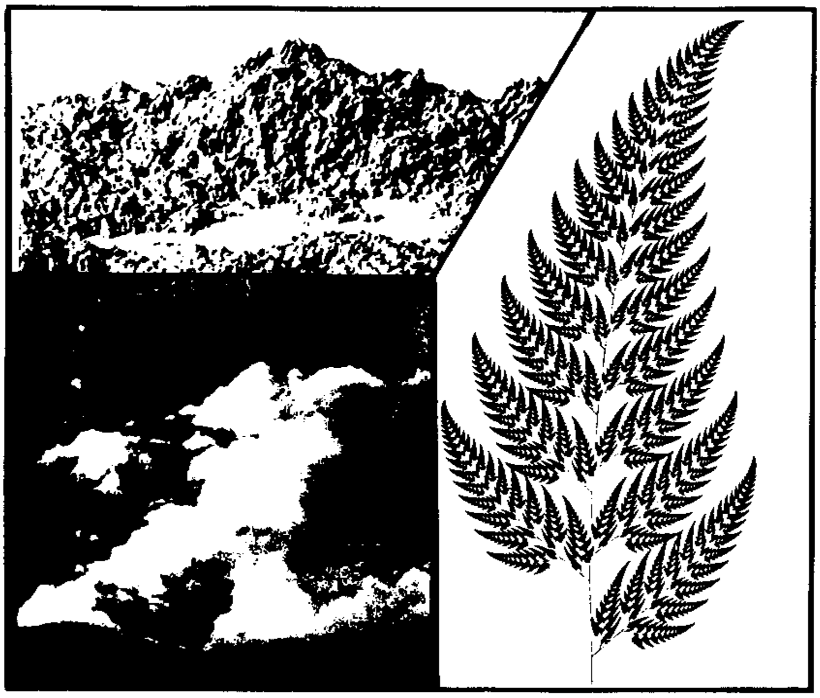
Unfortunately, the Generalised Relativity theory and the Quantum theory are irreducible one to the other. This fact has left the physician in a schizophrenic situation being able to properly account for the phenomenons of the microscopic world and the macroscopic world separately but not to unify its understanding of them.

We have a similar problem in media research.

At global level, one may legitimately use a concept of audience to refer to the readership of a magazine and media research provides us with adequate statistical tools to describe it. Similarly the media behaviour of a given individual can be modelled properly with usual probability techniques.
However it is technically difficult to integrate the individual and the global viewpoints in a unique theoretical framework. One may argue that when we are using a mediaplanning model, we consider the individual behaviour of the respondents and that the global performances are derived from them. Although, this is true it is instrumental not explanatory : the understanding of the global performances are not the consequence of the understanding of the individual behaviours.

In order to overcome the Quantum theory versus the Generalised Relativity theory dilemma, a cosmologist named Laurent NOTTALE has recently used the idea that the underlying phenomenons are fractal. In simple words this means that the logic of the phenomenons depends on the magnitude at which the exploration is made.

Fractals are strange geometric figures that reveal different coherent and repetitive structures depending on the magnitude of the observation. Here are some examples borrowed from B. MANDELBROT who first studied these mathematical entities.



A lot of things seems to be fractal in nature like forest and trees, mountains skyline, clouds... etc.
Using this approach Laurent NOTTALE has successfully unified, the four fundamental forces of the universe in a unique mathematical framework.

The proposal of this paper is to draw some preliminary lines towards a general purpose mathematical framework useful for media phenomenon description.

2- The Personal Probability Route

2.1. Mediaplanning

Since New Orleans, we have worked over the past Readership Symposiums towards a better mutual understanding of the different methodologies which are used in the world and more recently toward harmonisation. This is indeed essential and should be regarded as of a great value.

However we cannot wait any more for a global questionnaire harmonisation that is still a long way ahead despite of the real progress already made.

After twenty years of processing experience of many different audience surveys data my belief is that there is a sure way to make studies parallel taking as a fact of life that the methodologies are many. The benefit of it is obviously to use a common conceptual framework to describe the audience phenomenons.

In practice, most of the every day usage of media audience databases is mediaplanning. Almost, all (if not all) of the commonly available models can operate from a notion of individual probability of exposure.

Let p_i be the probability of exposure of respondent i

Let π_i be it's sample weight

By summing up the probabilities we get the audience of the measured vehicle :

$$A = \sum \pi_i p_i$$

Formula models use at most four notions.

The one and two issues coverage that can be computed by summing over the respondents :

$$C_1 = \sum \pi_i p_i$$
$$C_2 = \sum \pi_i (2p_i - p_i^2)$$

The maximum audience :

$$M = \sum \pi_i (p_i \neq 0)$$

The duplication between vehicle A and B :

$$D_{AB} = \sum \pi_i p_i^A p_i^B$$

Full binomial models and simulation models directly operate on the individual probabilities calculating a binomial or a markovian distribution for each respondent or group of respondent and aggregate vehicle by direct combination (convolution) of such laws.

Hence in all cases individual probabilities of exposure are a simple and efficient way to deliver audience data.

However handling individual probabilities used to require computer resources. But due to modern micro processors and sophisticated computing algorithm, this is no longer an issue.

2.2. Ascribing probabilities

Since audience research provides us with a fair statistical image of the population, it is common practice to ascribe to each respondent a probability of being exposed to an advertising vehicle.
Statistical techniques designed to do so have improved and are less demanding in terms of questionnaire length. We usually compute probabilities from magazine audience data by use of a technique that J.M AGOSTINI introduced in 1964. This technique goes like this.

Given a read/not read question for an average or specific issue and a declared habit of reading (numerical or verbal), each respondent within the same habit group is ascribed a probability equal to the observed proportion of issue readers within that group.

The weakness of this approach is that the number of different probabilities is small (as many as habit group) which has the nasty effect of inflating the coverage. Taking this fact into account some have chosen to do the computation within each target group. This is a wise thing to do as long as the target is not too small otherwise the probability estimate will have a huge variance.
A better solution is to get the probabilities by automatic segmentation as it has been the case in Germany for some years now.

We have personally worked along a different route which has the benefit to be applicable in a wide group of cases including those where we have no frequency question. The idea behind this method called probability by neighbourhood uses the fact that a probability can be defined as a proportion of yes (or win) within a set of similar cases.

First, each respondent is mapped onto a multidimensional space according to it's characteristics. This being done, we look for who are the closest neighbours of each respondent and we ascribe to this respondent the observed proportion of exposures within his vicinity group.

Here are the results of the three methods of probabilisation applied on data from the study measuring the French national newspapers.

Schedule with 19 insertions		**AGOSTINI**	**NEIGHBOURHOOD**	**SEGMENTATION**
REACH		20,8 %	21,5 %	19,5 %
QUANTILES	**1**			
	2	10,7 %	11,1 %	9,8 %
% of GRP	**3**	11,7 %	11,5 %	12,3 %
	4	24,7 %	24,8 %	24,1 %
		52,9 %	52,6 %	53,8 %

TARGET SIZE = 21 207

		NUMBER OF ISSUES					
COVERAGE BUILT UP		**1**	**2**	**3**	**5**	**10**	**20**
LE MONDE	- AGOSTINI	2.9	4.5	5.7	7.5	10.6	14.4
	- NEIGHBOURHOOD	2.9	4.6	5.9	7.7	11.0	15.1
	- SEGMENTATION	2.9	4.5	5.5	6.9	9.2	12.0
FRANCE SOIR	- AGOSTINI	1.4	2.2	2.8	3.7	5.4	7.6
	- NEIGHBOURHOOD	1.4	2.2	2.9	3.9	5.9	8.6
	- SEGMENTATION	1.4	2.2	2.7	3.4	4.6	6.3
L'EQUIPE	- AGOSTINI	3.4	5.5	7.0	9.3	13.1	17.4
	- NEIGHBOURHOOD	3.4	5.5	7.1	9.5	13.5	18.1
	- SEGMENTATION	3.4	5.4	6.8	9.0	12.5	16.2
LIBERATION	- AGOSTINI	1.6	2.5	3.2	4.2	5.8	7.9
	- NEIGHBOURHOOD	1.6	2.6	3.3	4.3	6.2	8.7
	- SEGMENTATION	1.6	2.4	3.0	3.8	5.2	6.9

You can see that the results show a slight decrease of coverage when using the segmentation method and no significant changes with the neighbourhood method although it uses no frequency question.

The conclusion of this test could be :

> *Use segmentation whenever you have a rich complete questionnaire and relay on the neighbourhood method whenever you have limited information.*

2.3. Individual behaviour

As mentioned earlier, the above process does not inform us on the real individual behaviour each respondent is not considered *per se* but is rather as a statistical atom used to described a global phenomenon. A similar ambiguity exists in physics : statistical mechanics provides a fair description of the phenomenons that occur in a gas by modelisation of the behaviour of the molecules in this gas but does not provides any description of the behaviour of a given molecule.

Then the question is :

are we able to derive a genuine model at individual level that will overall be coherent with the global one ?

The first thing is to introduce time and to define a notion of personal probability of reading a given issue at time t. Such models are called *stochastic models* their mathematics are complex and have been developed in more technical papers.

An important result is that it is possible to set up such models in such a way that their consequence at higher level lead to the usual full binomial model. Further work is still required to lead to more complex global models such as the Markovian one. This is to say that time probability of reading is a concept that resolve to personal probability of reading when we increase the magnitude of our observation from individual to global level. This is a kind of fractal embedding of two models, that can be compared to the fractal unification approach in physics.

At that point, I would like to emphasis a crucial point : the individual behaviour being described using a notion of probability of reading a given issue at time t, we admit the fact that the exact behaviour of an individual at time t (he or she read or do not read) is undecidable.

This fact is similar to the fact that according to Quantum theory one observer cannot know at the same time the location and the speed of an atomic particle.

We have to recognise that if we want to know perfectly and at any time, the exact behaviour of someone, we necessarily interfere with the phenomenon and this spoil the observation. The only thing that we get to know is a probability of this person behaviour.

This should remind us of some well raised critics against too intrusive methods of measurement.

2.4. From exposure to effect

The question is does the probabilistic framework allow us to go from opportunity to see and exposure to advertising effect.

I have recently performed some work which is of some interest for this question on IPSOS advertising post test data bank in co-operation with this French market research company team.

Our aim was to link at individual level the probability of exposure with advertising effectiveness indicators such as memorisation, appreciation and others. The nature of theses indicators is not the subject here nor the way they are collected.

In order to link the two types of probabilities, we have introduced a special form of linear model that generalise what is called logit models. The fit is fair to excellent in most cases.

This allows us to consider at the higher degree of magnitude of the full advertising campaign that the probabilistic route is still the way to go.

3- One Step Beyond

It is difficult to dare a conclusion on a subject such as the one which motivates my paper today, since its speculative nature should be considered as an invitation to follow-up the route and get more deeply into practical details.

However, I would like to raise a last comment.

We have seen how the probabilistic framework is adequate to model micro and macro audience phenomenons and how it can be used in practice. We have mentioned also recent work that take place at a higher degree of magnitude leading to new probabilistic models applicable to built more efficient advertising campaign.

My question is : is there a link back from the point that we have reached to individual purchasing behaviour ?

I have no answer so far but a possible hint.

The probabilistic models that we have used to relate probability of exposure and advertising effectiveness indicators come from a class of models that is also used by statisticians to model the probability of buying goods. Moreover, if we introduce time and brand loyalty phenomenons some researchers advocate to consider Non Stationary Individual Markovian Models. Last but not least individual product consumption behaviour requires stochastic models.

This is quite close to what we have used so far to model audience phenomenons following the Personal Probability Route.

My hint is :
let's go one step beyond on this route and may be we will be back home.