REVISITING MEDIAPLANNING MODELS ASSUMPTIONS

Gilles Santini, G.S. IT Services

1. INTRODUCTION

Addressing the growing demand of the advertisers for more accountable media decisions and increased vehicle selectivity the industry needs to revisit its current practice and decide whether the aging mediaplanning models in use are still valid or not.

Most of them have been designed at a time when limited processing capacities imposed simplified mathematical assumptions in order to make possible real time evaluation of the reach and frequency distribution of the media schedules.

Also, one could wonder whether it makes sense or not, to built sophisticated optimizers if the basic media planning models behind lack of realism.

The purpose of this paper is to revisit the nature of some of the assumptions underlying the current mediaplanning models and to identify new routes for improvement.

2. LIVING FOSSILS

Some media research pioneers attempted in the fifties to formulate mathematical law governing media audience accumulation under the form of functional relationships.

Politz magazine reach logarithmic law provides a well known example of such an attempt:

$$R_n = a + b\log(n) \tag{1}$$

Agostini's improved formula goes along the same lines of thoughts:

$$R_n = 1 - \prod_{k=1}^{n}\left(1 - \rho^{\log_2(n)} R_1\right) \tag{2}$$

The necessary parameters (a, b, ρ) where easily calculable from the one and two issues reach R_1 and R_2.
The above formulas or others with the same flavour have played an important role in the development of the mediaplanning expertise and although they may look naïve to us nowadays they where audacious breakthroughs.

They deserved to be mentioned as such but also because they are archetype for the belief shared at that time that mechanist laws were governing audience accumulation and more generally advertising effect.

Such a belief has faded away but one can wonder whether the so called "industry curves" or "ad hoc rules" still in use, which clearly are cousins of the above oldies are something else than living fossils which should not belong to the current mediaplanning practice.

3. PROBABILITIES MISCONCEPTIONS

It has been several decades since one started to use probabilities in media research.
However it is for sure the most frequent source of misunderstanding if not misconception.

In order to better understand the problems, it is useful to return to the notion of probability which can be defined in several ways among which two are relevant in media research.

Adopting the point of view of Laplace, one can define probability as,

- ***the ratio of the number of favorable cases over the number of possible cases***

and from an empiric point of view, to substitute to it, the limit value of the observed frequency of favorable cases, for an infinite number of trials.

More precisely if after n replicated trials one can observe π_n positive events the probability of a positive event is defined as:

$$p_L = \lim_{n \to \infty} \frac{\pi_n}{n} \tag{3}$$

This point of view is prevalent when one possesses a declaration or the observation of an individual rate of exposure to the different issues of the vehicles.

An alternative way is to base the notion of probability on that of similarity and to define the probability of an event by means of:

- ***the rate of realization observed in a domain of similarity of the event concerned***

More precisely if within a group of m similar individuals one can observe π_m positive cases the probability of a positive event is defined as:

$$p_S = \lim_{m \to \infty} \frac{\pi_m}{m} \tag{4}$$

This definition is well suited whenever one has an interest of collective nature for a group of individuals.

Clearly one cannot observe an infinite number of replicated trials nor study an infinite number of similar individuals so n and m will not be infinite but only big (whatever it means but this is just a technical problem which can be handled by statisticians).

If the probability of reading is homogeneous in time and within the similarity class the two definitions lead to the same probability estimation.

However the way in which the probabilities of reading p_L and p_S combine is different unless one accept a fairly strong hypothesis.

To understand this fact which is behind a great deal of misconceptions it is useful to conduct a controlled simulation.

Let us consider an artificial media world populated with virtual readers called *Lectrons*.

A is a weekly magazine . A_1 is the first issue of the year, A_2 the second one and since in that world the year is square with 12 months of 4 weeks each, A_{48} is the last issue of magazine A in the year.
Let us assume that each Lectron has the same probability p_0 of reading A and that such a probability does not change over time. So the probability of reading the k^{th} issue A_k is p_0 .

Clearly we expect to have:

$$p_L \;;\; p_S \;;\; p_0 \tag{5}$$

Following are the results for a population of 10 000 Lectrons reading a magazine with an average audience of 20% as simulated by a Monte Carlo type computer program.

m	p_m	n	q_n
1000	20.452	4	19.920
2000	20.279	8	19.884
3000	19.997	12	20.056
4000	20.017	16	20.102
5000	20.008	20	20.052
6000	19.986	24	20.061
7000	19.982	28	20.046
8000	19.974	32	20.126
9000	19.992	36	20.118
10000	**20.021**	40	20.068
		44	20.067
		48	**20.021**

Up to small fluctuations as expected the larger n and m the closer to p_0.

If we accept that the knowledge of the fact that Lectron i has read issue A_1 does not tell us anything about whether Lectron j has read or not the same issue or any other given one A_k we should believe that:

$$\Pr\{\, j \text{ read } A_k \mid \text{ knowing that } i \text{ read } A_1 \} = p_0 \tag{6}$$

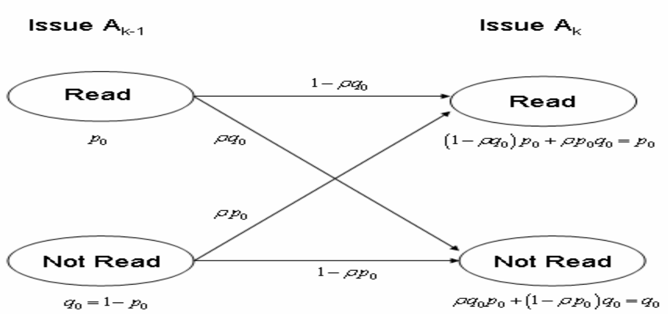
This is a rather mild hypothesis of independence that we should be ready to accept. In this case the probability to read both issues A_{k-1} and A_k is p_0^2.

On the contrary it is most certain that the past behaviour of reading of Lectron i tells us something about its future behaviour, so we should consider that:

$$\Pr\{\, i \text{ read } A_k \mid \text{ knowing that } i \text{ read } A_1 \} \neq p_0 \tag{7}$$

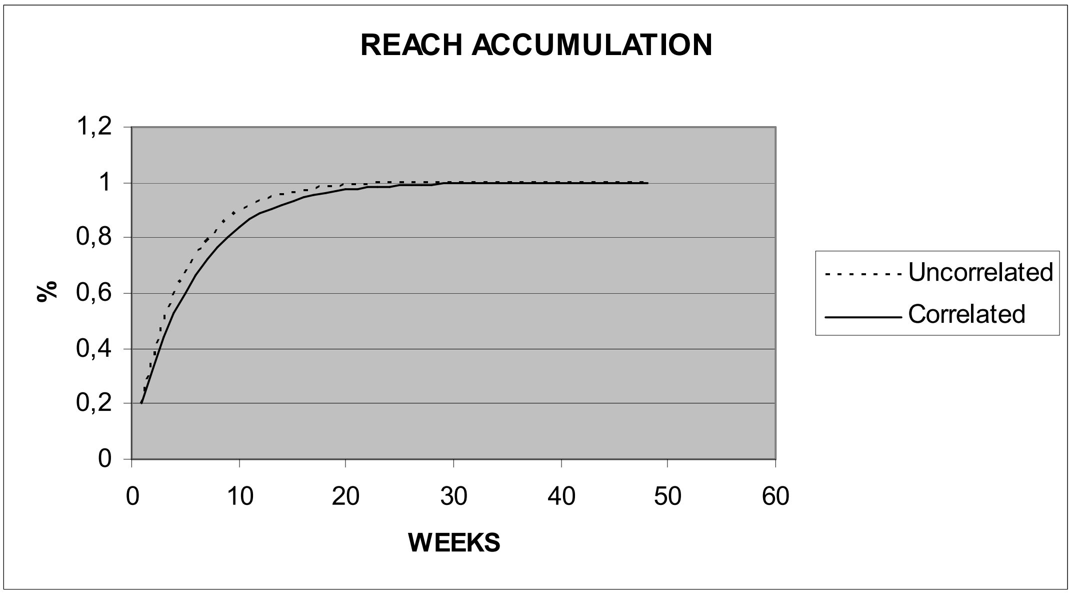
Let us assume further that one can draw as much knowledge on $\Pr\{\, i \text{ read } A_k \}$ from the fact that i has read A_{k-1} than from the observation of its total past behaviour.

This can be modelled by a Markovian process with two states:



Although the probability of reading issue A_k is stable and equal to p_0 the probability of reading both issues A_{k-1} and A_k is now equal to $\left(1-\rho q_0\right) p_0 = \rho p_0^2 + \left(1-\rho\right) p_0 \neq p_0^2$ as previously unless $\rho = 1$.

If we simulate numerically such a case we can produce the following graph which show the difference in reach accumulation using p_S or P_L i.e. uncorrelated or correlated readings.



It exemplifies the difference in nature between the two approaches.

This being said one should wonder why common mediaplanning model do not use p_L type probabilities and do not address the issues dependency phenomenon properly.

The problem is that RR technique applied to non identified issues can only provide similarity type probabilities such as p_S leading to a systematic bias in the reach calculation.

This major flaw is masqueraded as a benign defect by the following cover up.

4. THE INDEPENDENCE COVER UP

If instead of considering that every one in the population has an equal probability p_0 we could imagine that such a probability vary from an individual to another one with an overall mean equal to p_0 .

Let us first consider that the population splits into two groups of equal size one group with a probability of reading equal to $p_1 = p_0 - \alpha$ and the other one with a probability of reading equal to $p_2 = p_0 + \alpha$.

Let us assume that we use p_S type probabilities.

For each individual in the first group the probability of reading two issues of the magazine is p_1^2 and for each individual in the second group it is equal to p_2^2 .

For the total population the probability of reading an issue of the magazine is

$$\frac{p_1 + p_2}{2} = p_0 \tag{8}$$

while the probability of reading two issues is:

$$\frac{p_1^2 + p_2^2}{2} = \frac{\left(p_0 - \alpha\right)^2 + \left(p_0 + \alpha\right)^2}{2} = p_0^2 + \alpha^2 \neq p_0^2 \tag{9}$$

It looks as if the nasty independence assumption was not there but it is a mere cover up due to the stratification of the population.
Moreover departure from independence is always towards the same direction (i.e. positive correlation since the added term α^2 is positive.

If instead of considering the over simplistic case of two equal populations we let the probability of reading vary across all individuals in such a way that the fluctuations around p_0 vanishes out, it does not change the results because of the following algebraic fact:

$$\begin{aligned} &\frac{1}{I}\sum_{i=1}^{I} p_i^2 = \frac{1}{I}\sum_{i=1}^{I}\left(p_0 + \alpha_i\right)^2 = \frac{1}{I}\left[\sum_{i=1}^{I}\left(p_0^2 + 2p_0\alpha_i + \alpha_i^2\right)\right] = p_0^2 + \frac{2}{I}p_0\cancel{\sum_{i=1}^{I}\alpha_i} + \frac{1}{I}\sum_{i=1}^{I}\alpha_i^2 \\ &\Rightarrow \quad \frac{1}{I}\sum_{i=1}^{I} p_i^2 = p_0^2 + \frac{1}{I}\sum_{i=1}^{I}\alpha_i^2 \geq p_0^2 \end{aligned} \tag{10}$$

What one should remember and may be reconsidering is that at the core of the usual mediaplanning models sits the independence assumption of reading two given issues.

5. BETA-BIMOMIAL ROLLER-COASTER

Let us be practical and accept the fact that we can presently do nothing but go along with the independence of issues assumption.

So let p_i be the probability of reading a given magazine of i .By construction the average of the p_i is $\overline{p} = p_0$.

It is common practice to model the distribution of the individual probabilities by a statistical law called a Beta.

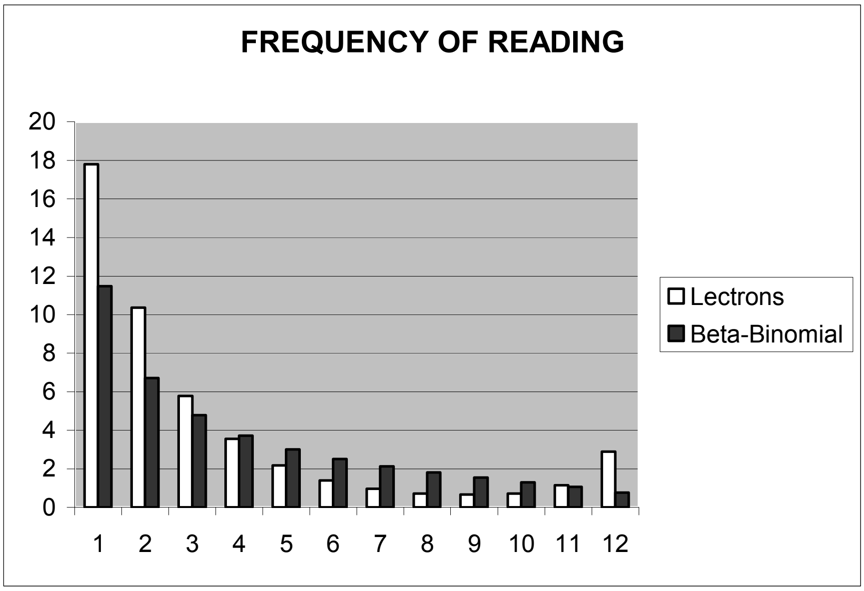
Beta are easy to work with and enjoy some sort of statistically legal status as being THE type of law to use with binomial outcomes as *read/not read*. The Bayesians, a weird tribe of statisticians, call Beta and Binomial conjugates laws. It is their combination which has given rise to Beta-Binomial mediaplanning models in the seventies of the past century.

The use of Beta and Beta-Binomial laws was not a bad idea then, when computers were a lot less powerful that my pocket camera but they lead to really bad effects whenever contrasted behaviour coexist in the population or when media fragmentation induces group of zero probability of reading.

As an example one can easily produce a Lectrons simulation that illustrates a circumstance among others when the Beta Binomial is incapable of accounting for the complexity of the frequency of reading.

In this particular but frequent one, the underlying population is two fold with a subgroup of rather regular readers and another one with more volatile habits:

The Beat-Binomial framework cannot catch the shape of the frequency distribution and heavily underestimate the 3 months reach.



Reach :
Lectrons=48% vs Beta-Binomial=41%

Looking at the previous results one should reach a stage of more mature enjoyments and quit with the up and down of the Beta-Binomial attraction.

6. DUPLICATION PUZZLE

Someone making from an audience data file a cross-tab of two magazines RR questions is expecting to get the averages audiences as margins and the duplication on the diagonal. Strangely enough he may or may not get the same duplication figure from a media schedule with one insertion in each magazine depending on the type of mediaplanning model he uses for that purpose.

This type of discrepancy is puzzling for the beginner but any aged media researcher does know that media planning is concerned with insertions that may be distinct in time and that the duplication of two issues may refer to non concomitant issues.

Still are we really sure that the probability to read the k^{th} issue of magazine A and the k^{th} issue of magazine B is equal to the probability to read the k^{th} issue of magazine A and the $k^{th}+1$ one of magazine B. How should contribute to the level of duplication of Time and Newsweek someone who systematically switch from one magazine to the other one from one week to the next : should it be 0 or ¼ ?

Even worse, if we use the RR question applied to any issue read in the last period what are we actually getting on the diagonal of the above cross-tab ? And in order to avoid any nervous breakdown I do not dare to speak of the effect of parallel reading on it!

To be honest we should recognize that we have no strong understanding about how the figure on the diagonal of that cross-tab relates to the schedule audience accumulation.

We can either shut our eyes on the question and ask the media model builders to do the necessary cosmetic in order to output that figure as the 2 frequency level of a 2 insertion media plan or simply ignore it and use (at individual level) the independence assumption between issues $A_{k'}$ and $B_{k''}$ just as we have used it between $A_{k'}$ and $A_{k''}$ or $B_{k'}$ and $B_{k''}$.

To the best of our knowledge no significant breakthrough has been made recently to understand this puzzle and produce a more adequate handling of the correlation of reading between magazines and develop more convincing models that would account for it.

7. CONCLUSION

The conclusion of this paper could be a program of research in 5 points:

1. Reconsider systematically what are the hypothesis behind all mediaplanning models and examine if they are justified or acceptable in case they are still necessary foe computation.
2. Drop Beta-Binomial like models in favour of individual probability models.
3. Produce more accurate estimates of the individual probabilities of reading.
4. Develop ways to estimate the correlation which exists between different issues readership.
5. Design new individual models that would model properly the duplication or readership between issues of two magazines.

At the price of some efforts toward reaching those goals the media research industry would produce more realistic and mode effective mediaplanning models.