Management & Marketing (2009) Vol. 4, No.1, pp. 65-76

AN ALGORITHMIC APPROACH FOR MODELLING CUSTOMER EXPECTATIONS

Nicolae POP, Adriana AGAPIE, Nicolae TEODORESCU
Academy of Economic Studies, Bucharest

Abstract. *The scope of this article is to discuss the dynamics of formatting customer expectations in financial services-under two models for assessing cumulative learning in customer expectations. The first model is a classical Bayesian one, the second model is an entirely new application of the Repetitive Stochastic Guesstimation (RSG) algorithm.*
The traditional assumption of postulating that empirical data have been generated from an underlying probability has been questioned even by orthodox theorists. Our research strategy is to cast this problem in the form of an optimization problem and show that RSG algorithm will produce a relevant solution for the original economic problem.

Keywords: Bayesian updating, Computational economics, Customer expectations, Repetitive Stochastic Guesstimation.

1. Introduction

The scope of this article is to discuss, with the scope of measuring customer perceived quality, the dynamics of customer relationships under two models for assessing cumulative learning in customer expectations. The first model is a Bayesian one, the second is emanating from the philosophy of computational economics and it is an entirely new application of the Repetitive Stochastic Guesstimation (RSG) algorithm(Charemza, 2002).

Under both these two models are challenged some rarely questioned platitudes in the practice regarding quality and customer satisfaction, like:

„*It is necessary to exceed customer expectations*"

„If *a customer expects a bad level of quality and receives it, he/she will reduce his/her level of preference for the brand*"

„*Given two-equally priced options, the customer will choose the one with the higher expected quality*"

„*Management should always focus on its most loyal customers*"

Common sense conclusions in the recent scientific literature is that these assertions are too simplistic and management must adopt a dynamically, complex view regarding updating customer expectations. Question is, are the managers supposed to look at a probabilistic calculus or they can be given an algorithm together with some instructions, in order to asses by themselves characteristics of their clients?
In order to perform a comparison, the model of Bayesian updating expectations as presented in (Rust, Inman et. Al, 1999) will be reviewed.

2. Review of the Bayesian scheme for updating expectations

2.1. Description and formulas

Below it will be described the model adopted by Rust, Inman et. al (1999).

Basically, the customer has a prior distribution-exactly a normal one-regarding the average quality for the product or service.

Let denote the customer's prior distribution of the average quality *X* of a certain brand with $\pi(x) \in N(\mu, \tau)$ with μ being the a-priori expected mean and the variance τ^2 reflecting the customer's degree of uncertainty.

Perceived quality of a particular transaction (*Y*) is also considered to be modeled by a normal distribution.

The density function representing the customers' perceived quality of a transaction is denoted with $f(y) \in N(Q, \sigma^2)$, representing random variation arising from both variability of quality and errors in perception.

The joint distribution is $h(x,y) = \frac{1}{\sqrt{2\pi\tau\sigma}} e^{-\frac{1}{2}[(\frac{x-\mu}{\tau})^2 + (\frac{y-Q}{\sigma})^2]}$

From those two is deduced the a-posteriori distribution for the perceived quality of the transaction under study, as being the perception of an experienced customer-the first layer.

The predictive (marginal distribution) of *X* is calculated as $p(x) = \int_{-\infty}^{x} h(x,y)dy = \frac{1}{\sqrt{2\pi\rho_1}} \frac{1}{\tau\sigma} e^{-\left[\frac{1}{2}\frac{(x-\mu)^2}{\tau^2+\sigma^2}\right]}$ where $\rho_1 = \frac{\tau^2 + \sigma^2}{\tau^2\sigma^2}$, so if one scales the units so that $\tau^2 + \sigma^2 = 1$ then $X \in N(\mu, 1)$.

Then a second transaction is considered, whose perceived quality is also normally distributed-of mean the so-called *disconfirmation*, empirically determined-and given unchanged initial customer expectations is derived the a-posteriori distribution for the perceived quality of the transaction.

More specific, if some level of a quality, y_t, is observed on the next transaction, the disconfirmation is $\Delta_t = y_t - \mu$. The posterior distribution of *X*, $\pi(x / y_t) = \frac{h(x_t, y)}{f(x_t)}$ is calculated as $\pi(x / y_t) = \frac{1}{\sqrt{2\pi\tau\sigma}} e^{-\frac{1}{2\tau^2\sigma^2}[x - \tau^2\sigma^2(\frac{\mu}{\tau^2} + \frac{y_t}{\sigma^2})]^2}$ which is normal with mean $\mu + \Delta\tau^2$ and standard deviation $\tau\sigma$.

Thus, for instance, the posterior mean increases when disconfirmation is positive and the customer's uncertainty decreases, regardless of the outcome.

These steps are re-iterated once again and the next conclusion arises: expectations and future predictions move in the direction of the perceived level of quality; variance is reduced as makes sense from having more experience.

The predictive density of the quality of the next transaction, y_{t+1} given observed quality level y_t is $p(y_{t+1}/y_t) = \int_{-\infty}^{x} f(y_{t+1}/x)\pi(x/y_t)dy$ where $f(y_{t+1}/x)$ is normal with mean μ

and variance σ^2. This predictive density is shown to be normal with mean $\mu + \Delta\tau^2$ and variance $\sigma^2 + \sigma^2\tau^2$, reflecting the greater certainty created by experience.

Thus if the result of successive transactions combined with an initial expectation amended by an equal number of disconfirmations is in the form of an a-posteriori predictive distribution p(x), obviously normal, and if an exponential utility is considered (for its good mathematical properties –continuous, twice differentiable-)with good economic interpretations-expected utility is V(x)=$\int U(x)p(x)dx$ and takes a linear form on expected performance of a certain brand and variance measuring perceived risk or uncertainty of the brand in question. Lastly, the brand with the highest expected utility is estimated by running a multinomial logit model among the several brands.

2.2. Objections

Objections rose in this paper, attempting to model the dynamics of customer expectations, fall into two categories: *major* and *minor*.

Major: clients are assumed to have an initial expectation which is only altered by confrontation with successive transactions. Thus, the effect of personal learning and evolution is entirely deduced as a result of the confrontation with a single brand. After exposing the client- with the same initial expectation all time-to several successive brands, corresponding expected utility are estimated and the one with the highest probability is selected. We rather think that the client is exposed directly to several brands and his priors are affected by these, at the first stage, and after that, given an evolution to a superior understanding, clients' priors are changing in shape, not only in parameters. It is true that rationality, in the sense of adaptive processes is not something that is expected to exist all the time in economy. Once some knowledge seeded in consumers minds the need for continual marginal learning and the effects of cumulative learning might be properly modeled by an approach like Rust's. We argue that switching brands might be neither the result of an increase-in the sense of accumulation of the distinct variances nor the result of some maximization utility principle constantly done by the rational consumers

Minor: Allover, for the feasibility of calculus, normal distribution is assumed. This assumption is not only making the calculus easier, it also guarantees a final result of the same shape, with different parameters. Normal distribution, if metaphorically speaking, assumes that a certain conviction is coming from an infinite number of trials, identical and independent. Yet, if one has enough time to learn, why would like to stay on the same level eventually with another parameters and not change to a superior, next organization? On the other hand, if priors are coming in a certain form, they become convictions in relatively short time, assuming limited resources for learning.

In the same line, if for some customers prior information in terms of normal distribution can prove valuable, for some other persons this can be different. In a research situation this is just another way of stating that different customers may have different views. Also, it is possible that two customers working with the same model and prior information can arrive at different posterior beliefs if they base their prior information on different bodies of past data.

Our aim is to reconcile with dynamic models of customer satisfaction according to which perceived quality (or expected quality) is a single point- on the field of computational economics. Economists have long understood that the economy is a complex system. In complex systems agents residing on one scale start producing behavior that lays one scale above them. The traditional assumption of postulating that empirical data have been generated from an underlying probability space in which rational agents reside and economic institutions are located has been questioned with increasing vigor by orthodox theorists, even of the dominant school. We track the problem of evolving customers- behavior and choices -in a theory attempting to allow inexplicable emergence. Emergence is understood to be a process that leads to the appearance of structure not directly described by the defining constraints and instantaneous forces that control a system. The research strategy to tackle problems that are not fully understood is to cast them in the form of an optimization problem and hope that algorithms for its solution will produce a conceptual framework that is relevant for the original economic problem.

2.3. Repetitive Stochastic Guesstimation Algorithm

In 2002 it was published in *Journal of Forecasting* a paper containing the description of a stochastic algorithm –*Repetitive Stochastic Algorithm*, authored by Professor Woitjeck Charemza, specialized in macroeconomics, finance and econometrics, targeted for Eastern European countries. Unlike similar time series in well-developed countries, in countries like Poland, Romania, Bulgaria, Latvia macroeconomic time series were non-stationary and extremely short. Therefore, traditional econometric methods proved unsatisfactory. On the other hand, this professor was holding particular knowledge and intuition about the political processes in Poland and also was in close contact with representative scientific figures and politicians not only in Poland but also in Romania, Bulgaria and Latvia. With an impressive mathematical knowledge, sensing the right model to be applied and having preconceptions about the value of the estimated coefficients, he found himself cornered by the results delivered in traditional econometric framework. (The same was the situation with representative econometricians in Romania.) So he did the following: given a time series, a model (chosen to fit the evolution of the time series) whose coefficients need to be estimated, given some initial opinions about these coefficients' values together with some intervals in which these are supposed to lie, you first evaluate the error (the difference between the model, with guessed coefficients and the available data). Then you start to generate uniformly inside the initial intervals some 'potentially better' coefficients. Two criterions are used to pick the next better coefficients: the overall error and the distance between the data and the model weighted with some learning function. (Don't worry for now about this learning function!). Once better values are determined for the coefficients to be estimated, the intervals around them are narrowed and previous steps are reiterated. By decreasing the intervals' length, the stopping criterion is assured.

What comes first in someone's (normative) mind? If the initial values for the coefficients are assimilated with some means coming from a normal distribution, and if the length of the intervals is associated with some variance, then must be something resembling with the Bayesian calculus for updating expectations and further comparisons must be performed. Since a prior normal and present normal gives a posteriori normal, a preliminary comparison seems a feasible task. (This has been done in the original paper, and independent experiments with the similar results were obtained by me.) If coefficients to be estimated are assumed to be borne by an unknown distribution whose only mean and something similar to variance is revealed, then if the process of successive trials is the embodiment of some probability distributions, some final estimations can be available after a process resembling with the Bayesian updating expectations framework. (Eventually one can see an advantage in this algorithm since no theoretical derivations depending on the particular form of the distributions are needed.)

A second thought is that this algorithm resembles with some intelligent techniques, like genetic algorithms or simulated annealing, or neural networks. Why? Since it has to do with a vaguely learning process and with randomly generating potential solutions inside some intervals. Such comparisons were done both from a theoretical point of view and also on particular problems.

Lastly, in this framework, came the question about mathematical convergence of this algorithm. What has been achieved so far consists of two things and a conclusion.

When a large number of iterations and replications is performed, the learning function, like it appears in the original version, is useless.

Depending on the particular problem in hand it is possible to refine the search inside the intervals, this is instead of uniformly generate, sample from a particular distribution, so that the convergence of the algorithm is achieved. Two examples are available: a linear model (call it regression it you have 3 coefficients and 10 data or general linear model if you have 10 coefficients and 3 data) and a GARCH(1,1) model (where you can estimate the parameters based on only two observations).

Conclusion is that this algorithm together with some proofs for convergence can be assigned to the bounded rationality concept. This is due to the fact that ad infinitum is capable to find the correct solution and in short time is finding better solutions than initial ones.

In short, the RSG can be presented as follows:

Procedure Repetitive Stochastic Guesstimation

1. Set the iteration index to zero: j=0
2. Choose some initial values and intervals for the parameters to be optimised
3. Choose/compute the initial value for the learning rate λ_0
4. Randomly generate (guess) a new candidate solution, inside the current intervals
5. Compare the candidate solution vs. the current one – w.r.t. both criterions – and decide: accept or reject
6. If accepted, it becomes the current solution; otherwise, keep the old one
7. Repeat 4-6, several times, until a better solution is obtained
8. j = j+1, decrease the learning rate, decrease the intervals' lengths and go to 4
9. Repeat 8, until some STOP conditions.

Actually, there are three points where RSG takes advantage on other stochastic algorithms:

1. At the initial stage, by making use of the prior beliefs concerning the parameters to be guessed - according to the economist's expertise and intuition.
2. By successively restricting the search space from iteration to another, providing an asymptotic convergence of the algorithm in some extreme point.
3. By using two objective functions, instead of one.

Let us briefly discuss these main features of RSG. The first seems to be restrictive for general optimization purposes, by limiting the application area of the method to a relatively well-known problem This assumption makes the comparison against evolutionary algorithms (like GA, Evolution Strategies or SA) somehow improper, as the last ones are commonly used in the so-called Black Box Optimization problems – where no information on the objective function is supporting the search. However, a basis for the comparison exists: many authors in the evolutionary computation field recommend the insertion of additional information into the initial population (of a GA, e.g.), whenever this information is available. As for the individual-based algorithms (SA, e.g.) the choice of initial values according to the user's expertise or intuition is welcomed and easy to implement.

A straight implication of the first point is that there is dependence between the initial points and intervals considered and the RSG estimates. This dependence will be explored in the following sections. On the other hand, these prior beliefs about the initial values of the parameters and intervals can be quantified and analyzed in a Bayesian framework.

The idea of running a searching algorithm from some expected valued for the parameters, according to experts intuition also appeared in Marcet's (Marcet, 1991) method of parameterized expectations.

If one keeps in mind this resemblance then he can continue with Sargent's interpretations and give to the RSG estimates the interpretation of approximate equilibrium points in the process of computing rational expectations equilibrium.

The second feature makes the connection to the heuristics of „learning algorithms", namely by retrieving the common sense expectation of '*increasing the guesstimator's confidence by narrowing the interval from which the parameters are to be guessed, as time goes on'*, (Charemza, 2002). This is definitely an important difference against GA, with its immutable searching space all over the algorithm. Theoretically speaking, the possibility of limiting the searching area from an iteration to another is specific to individual-based algorithms only (notice that RSG enters this category), thus not to population-based methods (like GA).

On the other hand, the brute technique is similar to a search technique called *'Fibonacci search'*, first developed in(Kiefer,1953). Fundamental differences are at the practical functions used for decreasing the successive intervals. The *Fibonacci search* is essentially deterministic and thus the intervals of uncertainty are governed by some difference equations. In the case of RSG algorithm, the intervals of uncertainty are probabilistic since they are dependent to a probabilistic-weighted objective function.

Regarding the two criterions, namely the *unweighted* and *weighted* objective functions, the (penalty) weights in the last one are normally distributed according to the difference between the currently guessed and the previous best guess, (Przechlewski, Strzala, 1996). This makes RSG a dynamical optimization method, by making the objective function time dependent. Again it seems that a deterministic correspondent of this idea can be found in (Marcet , Sargent, 1989). Against the class of least-squares learning technique, RSG has the great advantage of being very easy to manipulate and able to deal with the undersized sample size problems.

3. Description of a simple experiment and results

The experiment described in this paper is designed for a straightforward comparison with the model tested and presented in (Rust, Inman et al., 1999).

101 students enrolled in a Master degree with the Academy of Economic Studies were asked to fill a questionnaire with no obligations. The questions were about performances of three insurance companies, A, B and C regarding life insurances and private pensions. Data regarding the net earnings from deposits done in the account of life insurance were generated from normal distribution, as follows: net earnings for company A were sampled from N (7%, 2%), company B came from N(8%, 4%) and company C ' earnings were sampled from N(6.5 %, 1.5 %) (in total, 10 observations).

At the beginning, three observations regarding performances in years -3,-2 and -1 were presented. Then students were asked to asses a probability of choice for each of the three companies. Then 'the latest records' regarding performance (year 0) was showed to the students and they were asked again to assign a choice probability to each of these three companies. Students were asked to imagine that they were bound to stick with the previous company for three next years, while observing neat earnings for all three companies. After observing the next three values for companies A, B and C they were asked to assign some percentage to their preferences regarding these three companies. Then they were said to imagine that they are free of contracts, they observe the performance of the companies in year 8 and asked to choose a percentage reflecting their stickiness with some company. Then another two observations were made available, respectively and students were asked to make a choice (yet also in terms of percentage, not necessarily exclusive).

There are two major points which make the difference to the experiences described in (Rust, Inman et al., 1999). First is that the subjects were exposed to similar transactions generated by normal distributions, yet of different parameters.

Secondly, percentages representing confidence degrees were considered, instead of forcing a straight choice. These percentages can be assimilated with probabilities, yet the sum of these three is not necessarily one. Their mixture is an indication of the weights- as degrees of certainty- rather a straight choice bounded by the constraint of sum to be equal to 1. When these percentages are not summing to 1 one can have an indication of an incertitude regarding the choice to be done.

In order to better understand the comparison with the Bayesian framework and the approach proposed in this paper the organization of the questionnaire will be explained in detail.

3.1. The questionnaire

Given the net earnings of three anonymous companies providing life insurances policies in previous three years, indicate the degree to which you could chose a certain one.

(I)

Company A	Company B	Company C
Net earnings in previously three years: 6.5 %, 5.8 %, 5%	Net earnings in previously three years: 9.1 %, 5.8%, 11.5%	Net earnings in previously three years:3.3 %,6.1%, 6.8%
1.For sure (100 %)	1.For sure (100 %)	1.For sure (100 %)
2.Rather yes (75%)	2.Rather yes (75%)	2.Rather yes (75%)
3.Neither ….nor..(50%)	3.Neither ….nor..(50%)	3.Neither ….nor..(50%)
4.Less probable (25%)	4.Less probable (25%)	4.Less probable (25%)
5.Definitely no (0%)	5.Definitely no (0%)	5.Definitely no (0%)

Suppose that now we have to express an option among these three and you also have the next information about performances in the current period. Indicate the degree to which you chose a certain company.

(II)

Company A	Company B	Company C
Net earning: 10.1 %	Net earning: 10.31 %	Net earning: 7%
1.For sure (100 %)	1.For sure (100 %)	1.For sure (100 %)
2.Rather yes (75%)	2.Rather yes (75%)	2.Rather yes (75%)
3.Neither ….nor..(50%)	3.Neither ….nor..(50%)	3.Neither ….nor..(50%)
4.Less probable (25%)	4.Less probable (25%)	4.Less probable (25%)
5.Definitely no (0%)	5.Definitely no (0%)	5.Definitely no (0%)

In the next three years you are bound to your previous choice, if you had a pregnant choice or to any of the companies if you had equal preferences and you observe the yearly earnings. If you could chose to move your contract from one company to another, please indicate the degree to which you would chose one of the three companies: A, B or C.

(III)

Company A	Company B	Company C
Net earnings in previously three years: 6.8 %, 8 %, 8.71%	Net earnings in previously three years: 4.2%, 10%, 12%	Net earnings in previously three years: 6.7 %,7.4 %, 4.5 %
1.For sure (100 %)	1.For sure (100 %)	1.For sure (100 %)
2.Rather yes (75%)	2.Rather yes (75%)	2.Rather yes (75%)
3.Neither ….nor..(50%)	3.Neither ….nor..(50%)	3.Neither ….nor..(50%)
4.Less probable (25%)	4.Less probable (25%)	4.Less probable (25%)
5.Definitely no (0%)	5.Definitely no (0%)	5.Definitely no (0%)

Now, in the forth year, being aware of the net earnings just published, please indicate your choice for one of the three companies.

(IV)

Company A	Company B	Company C
Net earning: 9.2%	Net earning: 10.8 %	Net earning: 6.2 %
1.For sure (100 %)	1.For sure (100 %)	1.For sure (100 %)
2.Rather yes (75%)	2.Rather yes (75%)	2.Rather yes (75%)
3.Neither ….nor..(50%)	3.Neither ….nor..(50%)	3.Neither ….nor..(50%)
4.Less probable (25%)	4.Less probable (25%)	4.Less probable (25%)
5.Definitely no (0%)	5.Definitely no (0%)	5.Definitely no (0%)

After the next two years you have to make the definite choice among the three companies. Looking at their earning, which of these you would choose?

(V)

Company A	Company B	Company C
Net earnings: 6.6 %, 5.8%	Net earnings: 7.8 %, 5%	Net earnings: 5.1 %, 7.6 %
1.For sure (100 %)	1.For sure (100 %)	1.For sure (100 %)
2.Rather yes (75%)	2.Rather yes (75%)	2.Rather yes (75%)
3.Neither ….nor..(50%)	3.Neither ….nor..(50%)	3.Neither ….nor..(50%)
4.Less probable (25%)	4.Less probable (25%)	4.Less probable (25%)
5.Definitely no (0%)	5.Definitely no (0%)	5.Definitely no (0%)

Some possible results

(I): A_5, B_3, C_1 (II) A_2,B_2, C_2 (III)A_2, B_1, C_4 (IV) A_2,B_1,C_3 (V) A_4,B_5,C_1
(I): A_2, B_3, C_2 (II) A_2,B_2, C_3 (III)A_2, B_4, C_4 (IV) A_2,B_3,C_4 (V) A_2,B_3,C_3
(I): A_5, B_2, C_2 (II) A_2,B_1, C_4 (III)A_1, B_4, C_3 (IV) A_1,B_2,C_5 (V) A_2,B_3,C_2

In contrast to traditional analysis we do not have (at the beginning) preconceived beliefs regarding participants' ability to guess 'the best' company, therefore the subjects are not counted and split into groups, rather their particular choice trajectories are separately analyzed. We are interested to what degree a person can make the difference among similar (in this particular case, normal) distributions with close means and different variances. We intend to compare each option with the Bayesian counterpart, assuming a bit more general situation, namely the subject being exposed to transactions emanating from normal distribution with different parameters. In parallel with this comparison we will indicate a scheme for modeling this problem with RSG and some numerical results.

Let's look at the choices for the next subject:
(I): A_5, B_3, C_1 (II) A_2,B_2, C_2 (III)A_2, B_1, C_4 (IV) A_2,B_1,C_3 (V) A_4,B_5,C_1

By looking at the first set, ((I): A_5, B_3, C_1) the subject is made aware of the range of variation for the net earnings and feel the preferred mean. In this particular case, the preferred is company C, and we can assume that the initial mean μ_0 is say, the average of the values 3.3%, 6.1 %, 6.8%, this is μ_0=5.4 % and the range of variation is [min (6.5, 5.8, 5, 9.1,5.8,11.5,3.3, 6.1,6.8) ; max(6.5, 5.8, 5, 9.1,5.8,11.5,3.3, 6.1,6.8)] which is for this example [3.3, 11.5]. In the second set ((II) A_2,B_2, C_2) the subject has to adjust his mean according to disconfirmations. Since these are coming from different distributions, even if they look somehow similar, we rely on the choice probabilities (weights) indicated by the individual. This is: probability of choosing A is p_1=0.75, probability of choosing B is p_2=0.75 and probability of choosing C is p_3=0.75.In order to 'guess' the subject's a-posteriori mean after this disconfirmation we search for μ_1 and σ_1 so that

$$p_1 10.1+p_2\, 10.31+p_3\, 7 - \mu - \sigma\ \varepsilon \qquad (1)$$

is at minimum, where $\varepsilon \in N(0,1)$

The third stage is about re-thinking the choice, so we imagine the subject is re-adjusting his mean and try to estimate closer the spread around that.
The weights associated by this individual to the three companies are now
*p*1=0.75, *p*2=1, *p*3=0.75. (Please note that this is not necessarily indicating a straight preference among the three companies.)

Therefore we search for $\mu_2 \in V_{\mu_1}$ and $\sigma_2 \in V_{\sigma_1}$ so that

$$p_1[(6.8-\mu-\sigma\varepsilon)^2+(8-\mu-\sigma\varepsilon)^2+(8.71-\mu-\sigma\varepsilon)^2]+p_2[(4.2-\mu-\sigma\varepsilon)^2+(10-\mu-\sigma\varepsilon)^2+(12-\mu-\sigma\varepsilon)^2]+p_3[(6.7-\mu-\sigma\varepsilon)^2+(7.4-\mu-\sigma\varepsilon)^2+(4.5-\mu-\sigma\varepsilon)^2] \quad (2)$$

is at minimum, where $\varepsilon \in N(0,1)$.

The forth stage is about choice, centered around the re-estimated preference, this is around μ_2 and not around the previous choice mean, μ_1.

In this forth stage p_1=0.72, p_2=1, p_3=0.5

One has to run the algorithm once again for determining μ_3 and σ_3 so that

$$p_1[9.2-\mu_3-\sigma_3\varepsilon]^2+p_2[10.8-\mu_3-\sigma_3\varepsilon]^2+p_3[6.2-\mu_3-\sigma_3\varepsilon]^2 \quad (3)$$

is at minimum and then in the last step, find p_1, p_2 p_3 so that

$$p_1[(6.3-\mu_3-\sigma_3\varepsilon)^2+(5.8-\mu_3-\sigma_3\varepsilon)^2]+p_2[(7.8-\mu_3-\sigma_3\varepsilon)^2+(5-\mu_3-\sigma_3\varepsilon)^2]+p_3[(5.1-\mu_3-\sigma_3\varepsilon)^2+(7.6-\mu_3-\sigma_3\varepsilon)^2] \quad (4)$$

is at minimum.

3.2. Preliminary results

Correspondent estimates for the parameters in equations (1) to (4) were estimated according to the sufficient conditions presented in (Agapie, 2008). Therefore the estimates are not dependent on the sample size.

For the $\varepsilon \in N(0,1)$ we did proceed to the next data generating process (DGP): we sampled from the standard normal distribution 100 realizations of $\varepsilon \in N(0,1)$ and for each of these were obtained estimators for the parameters involved. The results deliver averages for these estimators, over the 100 Monte Carlo simulations and their unbiasdness and complete convergence is ensured as presented in (Agapie, 2008).

Therefore, averaged values are presented below. As one can see, in this paper were calibrated the successive objective functions for the RSG in order that results match the customers trajectories on three different situations. This is the main achievement of this paper.

For the next trajectory, (I): A_5, B_3, C_1 (II) A_2,B_2, C_2 (III)A_2, B_1, C_4 (IV) A_2,B_1,C_3 (V) A_4,B_5,C_1, where the subject had an initial strong preference for Company A and ended with a strong preference for the Company C, having the smallest variance, RSG's estimates for the choice probabilities are consistent:

p_1=0.41, p_2=0.18, p_3=0.87

Similarly, for the second analyzed trajectory, (I): A_2, B_3, C_2 (II) A_2,B_2, C_3 (III)A_2, B_4,C_4 (IV) A_2,B_3,C_4 (V) A_2,B_3,C_3, where the subject had a mild preference for Company B and ended with a mild preference for the Company A, the RSG's estimates for the choice probabilities are again consistent:

p_1=0.68, p_2=0.47, p_3=0.29.

Lastly, for the third analyzed trajectory, (I): A_5, B_2, C_2 (II) A_2,B_1, C_4 (III)A_1, B_4, C_3 (IV) A_1,B_2,C_5 (V) A_2,B_3,C_2 where the subject had equal preferences for Company Band C ,he ended with a mild preference for the Company B,and once again the RSG's estimates for the choice probabilities are consistent:

p_1=0.69, p_2=0.32, p_3=0.87.

A further goal is to run this choice of objective functions on the rest of data, classify in clusters similar results and analyze their composition, eventually from the classical points of views: gender, age, education, etc.

4. Conclusions and further research directions

This paper is as attempt to show that in the particular framework of forming expectations customers can be considered problem solvers. RSG is playing thus the role of an algorithmic agent in the context of bounded rationality for inferring and learning, and further research will be devoted to empirical analyzes of some behavioral experiments. The idea will be in the end to get rid from any reliance on probability considerations and move this kind of problem in the field of algorithmic complexity theory.

References

Agapie, A. (2008) Convergence of the guesstimation algorithm. forthcoming in *Communications in Statistics: Theory and Methods.*

Charemza, W.W. (2002) Guesstimation, *Journal of Forecasting*, **21** , pp. 417-433

Kiefer, J. (1953) Sequential Minmax Search for a Maximum. *Proceedings of the American Mathematical Society*, **4**, pp. 502-506.

Marcet, A. (1991) Solving non-linear models by parameterizing expectations, *Mimeo, Northwestern University*.

Marcet, A., Sargent, T. (1989) Convergence of least squares learning with hidden state variables and private information, *Journal of Political Economy*, **97**(6).

Rudolph, G. (1997) *Convergence properties of evolutionary algorithms*, Verlag Dr. Kovac', Hamburg, Germany.

Rust, R. T., Inman, J. J., Jia J., Zahornik, A. (1999) What you don't know about customer perceived quality : The role of customer expectations distributions *Marketing Science*, **18** (1), pp. 77-92.

Przechlewski, T., Strzala, K. (1996). Evaluation and comparison of the RSG and RSG-GA, *ACE Project Research Memoranda*, 17.

Sargent, T. (1995) *Bounded rationality in macroeconomics*, Oxford University Press.

Wolpert, D., Macready, W. (1997) No free lunch theorems for optimization, *IEEE Trans. on Evolutionary Computation*, **1**(1), pp. 67-82.