

Analyzing the role of non-seasonal discounts in consumer expenses for a small market environment

Elmira Kushta¹, Miftar Ramosaço², Dode Prenga³

¹Department of Mathematics, Faculty of Technical Sciences, University "Ismael Qemali", Vlore

²Department of Mathematics, Faculty of Technical Sciences, University "Ismael Qemali", Vlora.

³Computational Physics Group, Department of Physics, Faculty of Natural Sciences, University of Tirana.

Abstract. Here we consider the observed changes in the consumer expenses in a district of Albania following market sales discount. By calculating latent variable as the usefulness of the discounts offered we interpret the net expenses surplus as related to the actual marketing activities. Here we use a reduced set of variables related to the house holding as gender and age of principal buyers, and some few marketing aspects as telephone calls, whereas as response variable we choose the average days of market visits and expenses. In the area we investigated we obtained that the offered discounts promote rising expenses which in turn tells that the consumer behave as minimalist buyer as a responses of higher prices and not limited budgeted. This supports the idea that the consumers in this area are mostly rational. Next we see that the most accented effect is related to the average costume expenses that suggest the marketing activity related to this category of consumers.

Keywords: Probit model, consumer behavior, distributions.

I. Quantitative analyses for consumer behavior

Consumer behavior is an opinion formation process which is difficult to be studied quantitatively and rather complex as evidenced in [1]. In a formal approach they are assumed to act rationality in their decision making by optimizing some utility function, but this last cannot be measured directly. Meanwhile the assumption of rationality does not hold always which is clarified by behavioral theories as discussed in [2], [3] etc. Behavior seems to be too complex to be studied and analyzed by deterministic methods. Even so, researchers and scholars have outdone this difficulty by using statistical tools and probabilities facilitating modeling in this case as presented in many textbooks. Mixed calculation using econometric optimization and network dynamics have been developed as for example in [4] and many other applications. Generally, the consumer's decision making process in buying is complex but econometrically known and measurable. Aside of general models and regressions, practical calculation have demonstrated their capacities to describe consumer behavior in specific systems as in [5], [6] and many others. In our recent work [7] we applied a logistic regression to identify the consumer profile in a specific area, the factors affecting their behavior and other parameters characterizing the system of consumer

attitudes and activities. Therein we have focused our attention in the fact that the stationary of the state should be considered in the framework of advanced analysis elsewhere proposed in [8] and [9] with mathematical reviews in [10]. In reference [11] a more advanced technique have been reported dealing with complexity in the behavioral models. From those and other references that we are not listing here, we acknowledged the importance of folded econometric and mathematical details for quantitative consideration for such systems. This study is intended to evaluate a marketing aspect as discounts for example, by specifically considering the nature of the state of the system, the possible presence of not-apparent factors as latent effects or hidden variables etc.

II. Predictor factors and latent Variables

Following recent works on similar our systems as presented in [14] and [15] we performed the survey in local area in an Albanian district with 130,000 inhabitant gathering data for 1500 consumers for a period of two months of costume trade and another month after non-seasonal; sales where applied. From them, only 1200 result with full records and therefore we use this shortened database for analyses. The city has no specific characteristics in the sense of economical level, ethnic cultures etc., one could assume with no doubt as practically

representative for all urban area of the country, hence no additional variables are expected to interfere in the system. Our data have few variables and we restrict Gender, Age group, market ticket holdings, contacts by phone, average normal and expenditures. As response we consider the dichotomous variable representing the effect of prices discounts on expenses similarly with the analysis in references [5], [6] in the framework of standard regression techniques. Here we assign the value 1 or 0 respectively if the average visits after sales is greater or not than the average visits before them, as observed in the equal time interval period. The same is applied for the average expenditures. From phenomenological point of view we have argued same expectation of how each variable could affect the dynamics and it represented in the Table 1 herein.

Table 1: Variables of the model

Epoch		Type	Values	Expectations
	Gender	categorical	1,2	Possibly differed effects
	Age	categorical	1,2,3,4	Possibly differed effects
before	Average visit on the market	categorical	[1-30]	informative variable
discounts	Total visits	nominal	[1-30]	informative variable
announce	Average expenses	nominal	Real Value	Depended variable
ments	Total expenses	nominal	Real Value	Depended variable
	Having Market Cards	binary	[1,0]	increase chance of contacts
	Having contacted by phone	binary	[1,0]	increase chance of contacts
after	Total Visits	nominal	[1-30]	Informative variable
discount	Average visits	nominal	[1-30]	Informative variable
announce	Average expenses	nominal	Real Value	Depended variable
ments	Total expenses	nominal	Real Value	Depended variable

The state of the system is characterized by average purchasing expenses, number of visits in the market. Therefore the aim of the analysis is focused in the evolution of those quantities if sales or discounts are applied. We realized it by analyzing the distribution of the response quantity measured in those two states respectively, and applying a probit model to estimate the utility underlining the dynamics observed.

III. Relaxation of the distribution after sales

We start from the inspecting statistical behavior of the system by analyzing the distribution for characteristic variables, and we observe that among many candidates the q-Gaussian mentioned before in our work and introduced by [8], [10] etc

$$p(x) = \alpha \left[\left(1 - \beta(1-q)(-\mu) \right)^2 \right]^{\frac{1}{1-q}} \quad (1)$$

fits the data better than other tested. In addition and to count for mixed multiplicative properties as usually expected for complex dynamics, we use q-lognormal as detailed theoretically in the reference [10]

$$p(x) = \alpha \frac{1}{x^q} \left[\left(1 - \beta(1-q) \left(\left(\frac{x^{1-q} - 1}{1-q} \right)^{1-q} - \mu \right) \right)^2 \right]^{\frac{1}{1-q}} \quad (2)$$

Q-distribution have been successfully used in and suggested in the studies for complex systems as generalized in the reference [10] and applied in [9] etc. Remember that q-additive and q-multiplicative process responsible for q-Gaussian and q-lognormal respectively, are defined by q-algebra as follows

$$a \oplus_q b = \begin{cases} a + b + (1-q)ab & a, b > 0 \\ 0 & a, b \leq 0 \end{cases} \quad (3)$$

$$a \otimes_q b = \begin{cases} [a^{1-q} + b^{1-q} - 1]^{\frac{1}{1-q}} & a, b > 0 \\ 0 & a, b \leq 0 \end{cases}$$

that clearly shows the q parameter is the measure of the distance from pure processes. If in (5) we denote a,b,c.. the probability for separate events, the probability of their occurrence becomes quite complicated as really happen in reality. If a process has two types of interactions say the additive and multiplicative properties, the distribution characterizing the state is more likely to be a q-Gaussian [6]. A q-lognormal will be considered if logarithm of variable is considered but the original equation in 5 does not shows directly the expected q-multiplicative behavior. Moreover the q-lognormal is a stable attractor only for q=1, therefore functions of type (3) are very sensitive toward q-parameter. For a correct use of such analysis we implement a careful bin optimization as described in [16]. After this short briefing with q-distribution applied, we expect identified the characteristic observables for the responses of consumer in the system.

Here we selected for preliminary analysis the number of consumers visits in the market and the average purchasing expenses. The number of visits is very important for the statistics because it report the overall reaction of the consumers without being limited form the budget constraints. The distribution of the expenses is by nature the most important econometric parameter. The average and the variance for parameters are measurable if the distribution is stationary so we can perform statistical analysis in stationary states or in those states where variance is finite. According to [8] this would happen when q-Gaussian parameter is q>5/3. In [10] the full alternative approach called q-Central Limit Theorem stated rigoristy of Next, we would like to know the attracting property of marketing activity which can be measured by the number of visits; therefore both cases have been analyzed form the stability point of view.

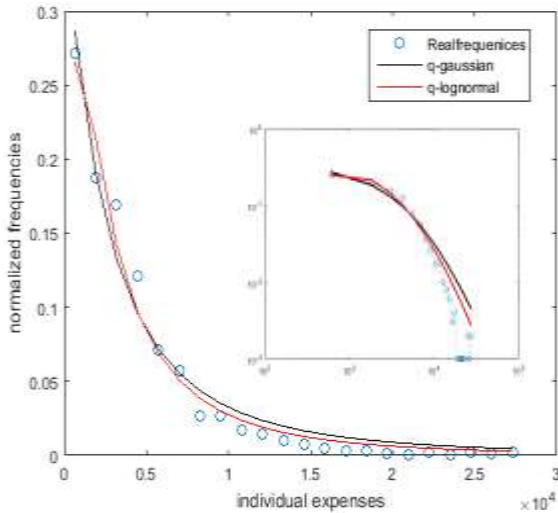


Figure 1: Distribution of visits average expenses. Normal trading period. Small picture shows log-log representation to better picture of the fit.

After realizing the fit to some expected common distributions, we observe that the parametric q-distribution had best statistics of fitting. The fitted curves are mostly q-lognormal within $\alpha=0.05$ restriction, whereas q-Gaussian has lower statistics, but is much less sensitive to the binning assize. Accepting functions of the type (2) as best fitted distributions, one can admit that the processes underlying the expenditures dynamics are q-multiplicative, hence very complicated. From the fitted q-lognormal we obtained the parameter $q=1.0001$ which report a nearly stationary lognormal if multiplicative processes are determinant. In particular it does not give the opportunity to measure the level of non-stationary as the difference $q-1=0.0001$ is too small. But in first equation of (3) we see that q-addition involve additive and multiplicative property, so for mixed processes it seems to be more significant. For this reason we prefer q-Gaussian for analysis of such behavior. Parameters q and adjusted R- Squared are [1.6531 0.9661] for q-Gaussian and [1.0001 0.9742] for the q-lognormal fitted. Therefore q-Gaussian tells that $q \sim 5/3$ that is in the boundary of definition for variances

$$\sigma_q = \frac{1}{(5-3q)\beta} \quad (4)$$

Next we considered the data for market visits and average expenditures after sales were applied. We obtain that the expenditure's distribution were found in a more stable states. The statistics for q-distributions fitted to the frequencies of consumer visits at the market again support the q-lognormal as best fitted function, but again by changing bin size we observe that q parameter in q-Gaussian changed only slowly whereas for q-lognormal it jumps from the value 1 with high margin. Therefore we consider

q-Gaussians for further analysis. Q-Parameters estimated and R2 for this case are found [1.6525 0.9778] for q-Gaussian and [1.0000 0.9974] for q-lognormal. We see that the stationary parameter q is nearly the same for the two series (before and after sales) but as we explained above the observation time for the second is much lower. So we accept that the state after discounts is more stable .

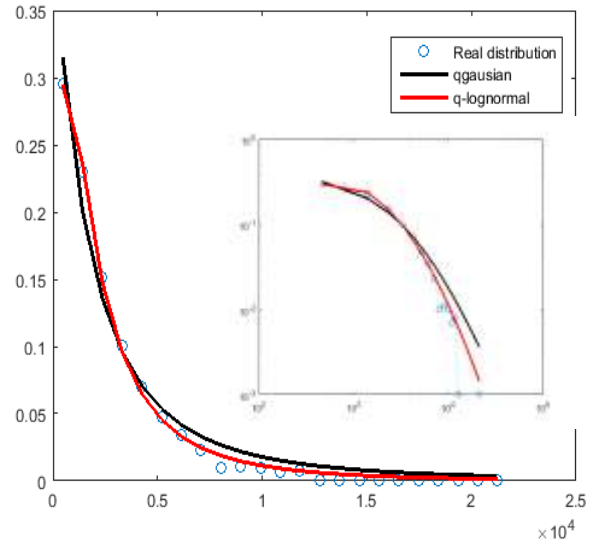


Figure 2: Distribution of average expenses (per visit) after trade discount announced.

We observe that announcing price discounts cause the number of visits to be higher and more averaged, that impose a more relaxed state which in turn could be related to a more stationary behavior or more common, typical consumer activities. In this sense we can this state for statistical analysis of the market, estimation of the representative values for expenditures and forecasting. We can simulate scenarios e.g., arrival of individuals with given set of properties using the distribution found here as PDF of the values for quantities discussed here. Moreover, we can approach this behavior as a response to a natural utility optimization and dominant factors to be typical ones, hence the disturbing terms could be neglected. Remembering that the time of post discount study has been as few as one third of the prior discount spanning period, the relaxation effect could be even more important and hence we can expect an intensive change of the probability of choice induced by discount announcement. This preliminary finding will support a more intensive change in cumulative probability CDF and therefore, the probit model could be more ascribing. In statistical aspect the result as of herein stated that the overall state is more stationary after applying discounts and we conclude that the first direct result of such marketing options in the relaxation of the system itself, and brining the distribution in the zone where the mean

and variance are measurable and well defined. In this sense, marketing studies and other analysis on the consumer behavior are likely to be more realistic in the period of sales. It is very important result for the market studies in the case of small market area as local districts and limited capacities for large inquiry that impose instability of the state of the system.

IV. Estimation of the utility and latent variables.

In the analysis of the effect of sales and discounts we proposed to analyze two scenarios: an hidden influences is assumed to be present or the outcome of the behavior is driven form a latent utility. In the first stage we perform factorial analysis to observe any possible reduction in the indicators, analyses as in [3]. We obtain that the number of latent variables could be larger than 2 but 95% of variance can be expressed in the tow first hidden components. Basing on the in the non-stationary of the state and the fact that after sales the state become more relaxed, we hypothesized that another hidden variable affect the purchasing behavior and checked it using MIMIC model. In this case we consider predictors variables in $X=[Gender, Age. Group, Average. Visits Average. Expenses Phone. Contact, Regular. Client]$, $Y=[Average Expenses after sales, Average Visits after Sales]$. Performing full structural model calculation under the assumption of an extra variable we obtain that for this choice, a single hidden variable will explain good the intermediary relationship between Causes X and consequences Y.

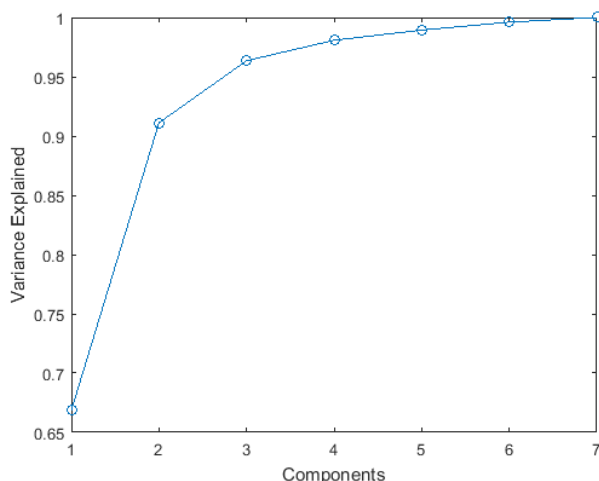


Figure 3: Identification of hidden or latent components

In table 2 we show parameter of linear equation $LV \sim Parameters * Factors$

Table 2: Parameters of hidden variables

Factors	Parameter Matrix A		Observed	Parameter Matrix B	
	Two HV	One H.V		SpendingAfterSales	1 HV-Mode
Free parameter	0.0046	0.526	0.0046	-7.9368	
Gender	1.0743	1.3994	1.0743	141.482	
AgeGroup	-0.2782	83.5887	-0.2782		
AverageVisits	-1.8217	-3.5924	-1.8217		
Average Spending	-0.3789	-0.4879	-0.3789		
RegularClient	0.0363	0.1666	0.0363		
TelephoneContact	141.482	0.0798	141.482		

Interestingly, the factors have different effects on average spending after sales and average visits. As seen in the Table 2, the parameters remain unchanged (up to 3 digits) in modeling with one and two hidden variable and therefore we restrict the model with one single hidden or latent variable. In this case the hidden variable could act as an interconnection between causes and outcomes. The reduction in the number of latent variables is plausible for the model because in this case we can use the utility as the intermediate variable ore stage in the consumer decision making. So far, we assume that the overall decision of the consumer to increase the expenses after discounts have been applied could be interpreted by a continues utility function

$$u_j = \beta_o + \sum_{i=1}^n \beta_i x_{i,j} \quad (6)$$

where j is the individual observation and i are variables. The response will be a dichotomous as follows

$$P(Y = y_j | X) = P(a_{i-1} \leq u < a_i | X) \quad (7)$$

and for our binary output there is only one point to be considered say the moment where the continuous probability take the value 0.5.

Firstly we consider the attractiveness of the discounts, so we examine increasing number of market visits after discounts were applied. Here we use chose as dependent variables the positive change in the average number of visits after discount; and for independent factors the gender of buyer, the age group and contacts by calls to announce the offers. By applying probit regression we observe that a good fit is obtained and the marginal errors are normally distributed as seen in the figure (3). The coefficients have been confirmed as different from zero within 90% confidence, whereas the free coefficient seems to not pass the test

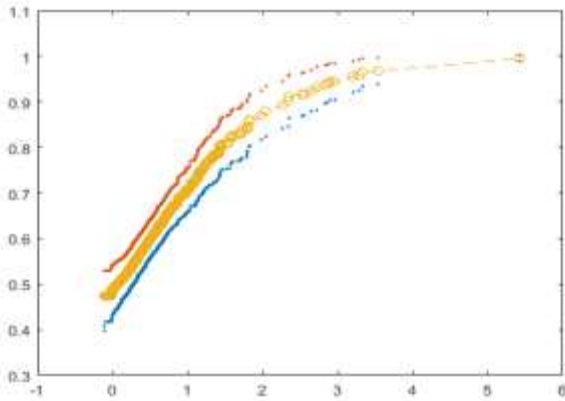


Figure 4: Probit regression for Increasing Expenditures after discounts

Therefore the utility of the attractiveness is obtained by probit regression as follows

$$Y^* = \{0.0698\} - 0.1367 * \text{GenderBuyer} + 0.0962 * \text{Age Group} + 0.0002 * \text{PhoneContact} + \varepsilon \quad (6)$$

From relation (3) we observe that the gender of buyers (F=1, M=2) is mostly decisive in the increasing number of visits in the market after sales, and usually male buyers are not more frequent in the market after discounts have been applied. The phone contact has a slight effect on it. The age group has comparable role to the gender of consumer. In Figure 4 is seen that the probability for more visits in the market is high for almost all the values of utility function (6) and only few values are less than 0.5. In this sense, for nearly all consumers' specifics, the marketing strategy (prices discounts) has been found attractive for peoples that respond by increasing the number visits in the market. Thus is the intermediate change on the consumer behavior. In the second stage, the final behavior is considered. Now the response variable is the change in expenditure measured by the natural function "is greater than" e.g., in absence of the marketing stimulus. It is possible that the buyer, under budget constraint, would respond to the discount spending the same quantity of money and therefore just buying some more goods, so this variable is meaningful and not trivially known. Here the independent variables include even average expenditures before sales and registered cards consumer. The first is expected to give information about which consumer category has increased the expenses, and the second could inform the role of being a formal consumer. Performing probit regression we obtain the utility function

$$y^* = -0.916 + 0.093 * C. \text{Gender} - (0.0133 * C. \text{AgeGroup}) + 0.0199 * \text{AverageVisit} + 0.854 \text{AverageExpenses} - 0.136 * \text{CardHolder} + (0.0013 \text{Call}) + 0.0396 \text{PeriodOpen} \quad (7)$$

In (7), the statistical significance is acceptable for all variables except Tel, Call and Age.Group (the

p.Value is high, ~0.3) so we put it in parentheses. Notice that their coefficients are small and so this does not affect the estimation of the utility so we kept them in the equation (7). Now we make use of the binary outcome expression using continues probability. The switching value for utility is

$$P(\text{ExpencesAfter} > \text{ExpencesBeforeX}) = P(u \geq 0.08) \quad (7)$$

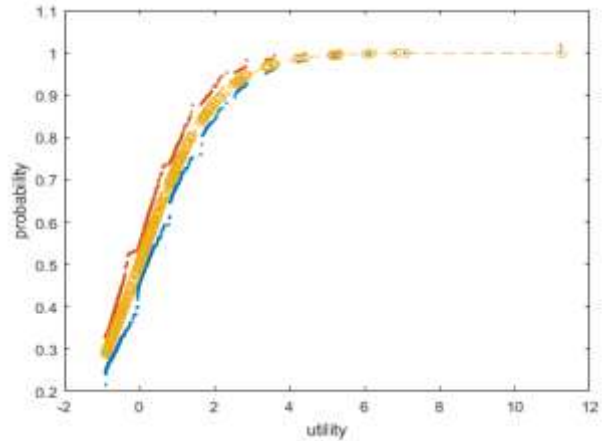


Figure 5: Full model probit regression

By using equation (6) we can realized conditions that (7) is fulfilled and therefore it is possible to forecast what happen with purchasing behavior for an individual with values

$$\text{Consumer}\{i\} \equiv n_1 n_2 \dots n_7 \quad (8)$$

This is done by just putting values (8) in equation (7). We observe that rational or cognitive issues weight more in the utility value as seen from the coefficients for female byer that usually behave as major house holdings buyers, common average expenses that indicate the level of budget in the purpose, the time of effectiveness for sales. The expected psychic parameters as telephone call have less effect in shifting the utility. By randomly selecting consumer (predictor) values according to the population considered, we see that the value of (7) is usually reached, therefore the conclusion herein seem to be a global tendency in the area studied. Remember that those findings are general characteristics because the distribution in this state has been acknowledged as being stationary.

V. Conclusions

We have realized an integrated analysis for consumer reaction toward marketing tactics and strategies which could be generalized methodically for larger area. We observe that "the state" of average expenses become more stationary after the discounts have been applied. Therefore, analyses of the market, measurements of quantities and statistical study for this system should be better performed on the after-discounts states. Particularly we conclude that the consumer reaction to the discounts was characterized by the increase of

spending itself, not only the volumes of items purchased. We identified the load of each factor in the increase of expenditures and acknowledge the utility form in this case. The deflection point on the decision is calculated using this utility. It give us the most probable feature of the system regardless of individual properties. Any combination of individual items that would produce the critical utility found, probably impose an increase on the expenses after temporally price reductions. From marketing point of view this work suggests that well studied and programed concessions will be beneficiary not only for the stock deleverage opportunity , but even to increase the profit itself. In the economic area studied herein, the most effective strategy should be based on the simulation for buyers of highest average expenditure, householdings oriented strategies. It seems that making calls or simulating consumers of specific age group is not much productive. The certainty of the results can be improved if this calculation could be extended for larger areas or specific industries which remain the focus of our future wroks.

VI. References

- [1] Claudio Castellano, Santo Fortunato, Vittorio Loreto.: Statistical physics of social dynamics. Rev. Mod. Phys. 81, 591-646 . April-June 2009
- [2] Jisana T. K. Consumer behaviour models: an overview. Volume 1, Issue 5 (May, 2014)
- [3] Jeff Bray. Consumer Behaviour Theory: Approaches and Models . <http://eprints.bournemouth.ac.uk>
- [4] N. J. McCullen, M. V. Ivanchenko, V. D. Shalfeev W. F. Gale. A Dynamical Model of Decision-Making Behavior in a Network of Consumers with Applications to Energy Choices. <http://www1.maths.leeds.ac.uk/>
- [5] Kadri G Yilmaz , Sedat Belbag. Prediction of Consumer Behavior Regarding Purchasing Remanufactured Products: A Logistics Regression Model. International Journal of Business and Social Research Volume 06, Issue 02, 2016
- [6] Hedeker, D. (2003). A mixed-effects multinomial logistic regression model. Statistics in Medicine, 22, 1433–1446.
- [7] E.Kushta, D.Prenga, F.Memaj. International Journal of scientific research and management (IJSRM). Vol 6 NO 03(2018).
- [8] Constantino Tsallis Computational applications of non-extensive statistical mechanics. Journal

- of Computational and Applied Mathematics 227 (2009) pp 51-58.
- [9] G.P. Pavlos, M.N. Xenakis, L.P. Karakatsanis, A.C. Iliopoulos, A.E.G. Pavlos D.V. Sarafopoulos. Universality of Tsallis Non-Extensive Statistics and Fractal Dynamics for Complex Systems. Chaotic Modeling and Simulation (CMSIM) 2: 395-447, 2012.
 - [10] Sabir Umarov, Constantino Tsallis, Murray Gell-Mann, Stanly Steinberg. Generalization of symmetric -stable Lévy distributions for $q > 1$. Journal of mathematical physics 51, 033502 2010
 - [11] Steiger, J.H. (1990), "Structural model evaluation and modification," Multivariate Behavioral Research, 25, 214-12.
 - [12] Sibiu Alma Mater University Journals. Series A. Economic Sciences – Volume 2, no. 4, December / 2009
 - [13] Horn, J. L. A rationale and test for the number of factors in factor analysis. Psychometrika, 30, 179-185.
 - [14] F. Mema. Marketing research in helping decision making” . “Ekonomia dhe Biznesi” Nr. 2 (22) 2006
 - [15] F. Mema. Statistical methods in acknowledgment of consumer behavior" International Review of Science, Innovation and New technology, Vol. 1, Nr 12, February, 2015
 - [16] Scott, David W. Multivariate Density Estimation and Visualization Papers Humboldt-Universitet Berlin, Center for Applied Statistics and Economics (CASE), no. 2004,16.

Author(s) Profile

Elmira Kushta, PhD student received the M.S. degrees in Mathematics from the Faculty of Natural Sciences, University of Tirana in 2010. She started as lecturer of statistics in the Faculty of Technical Sciences, University of Vlore. Now she is working on statistical and operational researches and simulation for consumer behavior and marketing.

Dode Prenga, Physicist, graduated in 1992 form the Faculty of Natural Sciences, is lecturer of Nonlinear Dynamics in the Department of Physics, Faculty of Natural Sciences, University of Tirana. He works in complex systems and interdisciplinary subjects.

Miftar Ramosaço, Dr, has graduated on 1989 in Mathematics from the Faculty of Natural Sciences, University of Tirana in 2010. She started as lecturer of statistics in the Faculty of Technical Sciences, University of Vlore. Now she is working on modeling and operational researches.