1-1-2009

Asymmetric Heteroskedasticity Models: A New Justification

Songsak Sriboonchitta

Vladik Kreinovich
University of Texas at El Paso, vladik@utep.edu

Comments:
Technical Report: UTEP-CS-09-02
To appear in *International Journal of Intelligent Technologies and Applied Statistics*

Asymmetric Heteroskedasticity Models: A New Justification

Songsak Sriboonchitta[1] and Vladik Kreinovich[2]

[1]Faculty of Economics
Chiang Mai University
Chiang Mai 50200 Thailand
email songsak@econ.cmu.ac.th

[2]Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
vladik@utep.edu

Abstract

Most existing econometric models such as ARCH(q) and GARCH(p,q) take into account heteroskedasticity (non-stationarity) of time series. However, the original ARCH(q) and GARCH(p,q) models do not take into account the asymmetry of the market's response to positive and to negative changes. Several heuristic modifications of ARCH(q) and GARCH(p,q) models have been proposed that take this asymmetry into account. These modifications turned out to be very adequate and efficient in describing the econometric time series. In this paper, we propose a justification of these heuristic modifications – and thus, an explanation of their empirical efficiency.

1 Introduction to the Problem

One of the main goals of econometrics. One of the main objectives of econometrics is to use the known values x_t, x_{t-1}, x_{t-2}, ..., of different economic characteristics x at different moments of time t, $t-1$, $t-2$, ..., to predict the future values x_{t+1}, x_{t+2}, ..., of these characteristics.

Comment. For a detailed description of econometric problems, ideas, and techniques, see, e.g., [4, 7].

First approximation: engineering models. A similar problem of analyzing time series x_t exists in engineering applications. So, historically the first econometric models simply used the formulas developed in engineering applications.

In engineering, most processes are stationary. It is known that stationary processes x_t can be well-described by auto-regression (AR) models:

$$x_t = a_0 + \sum_{i=1}^{q} a_i \cdot x_{t-i} + \varepsilon_t, \tag{1}$$

where ε_t are independent normally distributed random variables with 0 means and standard deviation σ – i.e., $\varepsilon_t = \sigma \cdot z_t$, where z_t is normally distributed with 0 means and standard deviation 1. The more terms we take in the AR model, i.e., the larger the value q, the better the corresponding AR(q) model describes the stationary process.

Heteroskedisticity (non-stationarity): a specific feature of econometric time series. In contrast to engineering time series, economic time series are usually non-stationary (*heteroskedastic*).

Specifically, in the economic time series, the empirical standard deviation σ of the remainder term ε_t depends on time. In other words, instead of a single value σ, at different moments of time t, we have different values σ_t. Thus, to appropriately describe the corresponding time series, we also need to know how this value σ_t changes with time.

Second approximation: models that take heteroskedisticity into account. The heteroskedisticity phenomenon was first taken into account by Engle [5] who proposed a linear regression model of the dependence of σ_t on the previous deviations:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \cdot \varepsilon_{t-i}^2. \tag{2}$$

This model is known as the Autoregression Conditional Heteroskedacity model, or ARCH(q), for short.

An even more accurate Generalized Autoregression Conditional Heteroskedacity model GARCH(p,q) was proposed in [2]. In this model, the new value σ_t^2 of the variance is determined not only by the previous values of the squared differences, but also by the previous values of the variance:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \cdot \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \cdot \sigma_{t-i}^2. \tag{3}$$

Several modifications of these models have been proposed. For example, Zakoian [12] proposed to use regression to predict the standard deviation instead of the variance:

$$\sigma_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \cdot |\varepsilon_{t-i}| + \sum_{i=1}^{p} \beta_i \cdot \sigma_{t-i}. \tag{4}$$

Nelson [8] proposed to take into account that the values of the variance must always be non-negative – while in most existing autoregression models, it is potentially possible to get negative predictions for σ_t^2. To avoid negative predictions, Nelson considers the regression for $\log \sigma_t^2$ instead of for σ_t:

$$\log \sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \cdot |\varepsilon_{t-i}| + \sum_{i=1}^{p} \beta_i \cdot \log \sigma_{t-i}^2. \tag{5}$$

Asymmetry: an additional feature of economic time series that needs to be taken into account. The above models such as ARCH(q) and GARCH(p,q) models are still not always fully adequate in describing the actual econometric time series. One of the main reasons for this fact is that these models do not take into account a clear *asymmetry* between the effects of positive shocks $\varepsilon_t > 0$ and negative shocks $\varepsilon_t < 0$.

It is therefore desirable to modify the ARCH(q) and GARCH(p,q) models by taking asymmetry into account.

Models that take asymmetry into account. Several modifications of the ARCH(q) and GARCH(p,q) models have been proposed to take asymmetry into account.

For example, Glosten et al. [6] proposed the following modification of the GARCH(p,q) model:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} (\alpha_i + \gamma_i \cdot I(\varepsilon_{t-i})) \cdot \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \cdot \sigma_{t-i}^2, \tag{6}$$

where $I(\varepsilon) = 0$ fir $\varepsilon \geq 0$ and $I(\varepsilon) = 1$ for $\varepsilon < 0$.

Similar modifications were proposed by Zakoian [12] and Nelson [8] for their models. The asymmetric version of Zakoian's model has the form

$$\sigma_t = \alpha_0 + \sum_{i=1}^{q} (\alpha_i^+ \cdot \varepsilon_{t-i}^+ + \alpha_i^- \cdot \varepsilon_{t-i}^-) + \sum_{i=1}^{p} \beta_i \cdot \sigma_{t-i}, \tag{7}$$

where:

- $\varepsilon^+ = \varepsilon$ for $\varepsilon > 0$ and $\varepsilon^+ = 0$ for $\varepsilon \leq 0$;
- $\varepsilon^- = \varepsilon$ for $\varepsilon < 0$ and $\varepsilon^- = 0$ for $\varepsilon \geq 0$.

The asymmetric version of Nelson's model has the form

$$\log \sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} (\alpha_i \cdot |\varepsilon_{t-i}| + \gamma_i \cdot \varepsilon_{t-i}) + \sum_{i=1}^{p} \beta_i \cdot \log \sigma_{t-i}^2. \tag{8}$$

Problem: the existing asymmetric modifications are very heuristic. The main problem with the existing asymmetric models is that they are very ad hoc. These model are obtained by simply replacing a symmetric expression ε^2 or $|\varepsilon|$ by an asymmetric one, without explaining why these specific asymmetric expressions have been selected.

The empirical fact that these asymmetric models work well in describing econometric time series shows that the selection of asymmetric function was indeed adequate. It is therefore desirable to provide an explanation for this adequacy.

Such an explanation is proposed in this paper. Specifically,

- we justify the general form of these expressions by the need for fast computations, and
- we justify the specific dependence on ε_{t-i} by the scale-invariance requirement.

2 Justification of the General Form of the Asymmetric Models

Objective: reminder. Our objective is to explain the existing asymmetric models for predicting σ_t based on the previous values ε_{t-i} and σ_{t-i}:

$$\sigma_t = f(\varepsilon_{t-1}, \ldots, \varepsilon_{t-q}, \sigma_{t-1}, \ldots, \sigma_{t-p}). \tag{9}$$

In this section, we justify the general form of the existing asymmetric models. A specific dependence on ε_{t-i} will be explained in the following section.

Need for fast computations. One of the specific features of econometric applications such as applications to trading is that the corresponding time series analysis needs to be performed as fast as possible.

Indeed, assume that a model predicts that a certain financial instrument will gain in price, and therefore, buying this instrument at the existing price can be profitable. The trader who is the first to predict this gain buys this instrument at the existing price. A buy – indicating an increased demand – usually increases the price of the instrument, so the traders who only later realized the potential gain have to pay more for this instrument – and thus, loses some of the potential profit.

The need for fast computations is well known in computational finances, leading use of very fast computers and fast algorithms.

Need for parallel computations. The fastest way to perform computations is to use the fact that several computers can work in parallel, performing different parts of the computations simultaneously and thus, reducing the overall time needed for computations.

Need for dividing the computation into simple pieces. For parallel computations, how can we decrease the computation time? In general, in parallel computations, we divide the computation into several pieces assigned to different processors. From the mathematical viewpoint, the resulting function is a composition of the functions computed by individual processors.

The simpler the function, the faster it can computed. Thus, to speed up computations, we need to divide the computations into pieces which are as simple as possible – and in which the total number of sequential steps is as small as possible.

First conclusion: importance of considering functions of one variable. How can we gauge the simplicity of a computation piece? First, it is reasonable to look at how much data we need to process. In general,

- the more data we need to process,
- the longer the corresponding computation will take.

Thus, to speed up computations, it is desirable to select computation pieces in which the amount of the processed data – i.e., the number of numerical inputs – is as small as possible.

The best possible case is when we have only one input, i.e., when we are computing the values of the function of a single variable.

Need for functions of several variables. Compositions of functions of a single variable are still functions of a single variable. So, to be able to process multiple data values, we also need to use some functions of several variables.

To speed up computations, it is therefore important to select the simplest possible (thus fastest-to-compute) functions of several variables.

Second conclusion: need to use linear functions. Functions can be

- linear and
- non-linear.

Clearly, linear functions are simpler and faster-to-compute than nonlinear ones. Thus, it is desirable to use linear functions of several variables.

Resulting computation scheme. As a result of the above analysis, we arrive at the following computations scheme. The computation consists of several sequential layers, in each of which:

- we either apply (in parallel) a function of one variable to each the inputs and/or intermediate computation results; such layers will be marked as NL (for Non-Linear),
- or we compute a linear combination of several inputs and/or intermediate computation results; such layers will be marked as L (for Linear).

These layers must intertwine. Indeed, if, e.g., we apply a function $y_i = f_i(x_i)$ of one variable to each input and then sequentially apply another function of one variable to the result $z_i = g_i(y_i)$, then we get the values $z_i = h_i(x_i)$, where $h_i = f_i \circ g_i$ is a composition of the functions f_i and g_i. We could have achieved the same result by simply applying the composition in a single step (i.e., in a single layer).

Similarly, if we first compute linear combinations and then again compute linear combinations of the computed linear combinations, then the results are simply linear combinations of the original inputs. These linear combinations could have been computed in a single step (i.e., in a single layer).

Thus, the sequence of layers must have the form L-NL-L-... or NL-L-NL-...

As we have mentioned, to speed up computations, we need to use the smallest possible number of layers, especially the smallest number of multi-input layers – i.e., layers of type L.

Smallest possible number of layers. The smallest possible number of L layers is 1.

In the simplest possible case when we only have a linear layer and no non-linear layers, then the expression (9) takes the form

$$\sigma_t = \alpha_0 + \sum_{i=1}^{q} \alpha \cdot \varepsilon_{t-i} + \sum_{i=1}^{p} \beta_i \cdot \sigma_{t-i}. \tag{10}$$

This expression is too simple and is nor adequate to describe asymmetric heteroskedasticity: since its dependence on ε_{t-i} is too asymmetric. Thus, in addition to an L layer, we need at least one non-linear layer. If we have a NL layer (in which a function $f(x)$ of one variable is applied) after the linear layer, then we get the dependence

$$\sigma_t = f\left(\alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i} + \sum_{i=1}^{p} \beta_i \cdot \sigma_{t-i}\right) \tag{11}$$

for some function $f(x)$. This expression also does not adequately describe the econometric time series, since the dependence on ε_{t-i} is too still too asymmetric.

When we have a NL layer (in which a function of one variable is applied to each input) *before* the L layer, we get a much more adequate dependence

$$\sigma_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \cdot f_i(\varepsilon_{t-i}) + \sum_{i=1}^{p} \beta_i \cdot g_i(\sigma_{t-i}) \tag{12}$$

for some functions $f_i(x)$ and $g_i(x)$. This formula already includes one actually used expression as a particular case: Zakoian's formulas (4) and (7). In Zakoian's formulas, $g_i(x) = x$, and the only non-linearity is in the expression $f_i(\varepsilon_{t-i})$ containing ε_{t-i}:

- the symmetric version of Zakoian's formula corresponds to $f_i(x) = |x|$; and

- the general (asymmetric) version of Zakoian's formula corresponds to $f_i(x) = x^+ + (\alpha_i^-/\alpha_i^-) \cdot x^-$.

If we allow NL layers both *before* and *after* the L layer, then we get a general scheme

$$\sigma_t = f\left(\alpha_0 + \sum_{i=1}^{q} \alpha_i \cdot f_i(\varepsilon_{t-i}) + \sum_{i=1}^{p} \beta_i \cdot g_i(\sigma_{t-i})\right) \quad (13)$$

that contains all the proposed formulas as particular cases. Indeed:

- ARCH(q) and GARCH(p,q) correspond to using $f(x) = \sqrt{x}$ and $f_i(x) = g_i(x) = x^2$.
- Glosten's formula corresponds to using $f_i(x) = (1 + (\gamma_i/\alpha_i) \cdot I(x)) \cdot x^2$.
- A symmetric version of the Nelson's formula corresponds to $f(x) = \sqrt{\exp(x)}$, $f_i(x) = |x|$, and $g_i(x) = \log x^2$; the general (asymmetric) version of this formula corresponds to $f_i(x) = |x| + (\gamma_i/\alpha_i) \cdot x$.

Conclusion. The need for fast computations explains a general form (13) that indeed includes all formulas proposed for describing asymmetric heteroskedasticity as particular cases.

3 Justification of the Specific Form of the Dependence on ε_{t-i}

Formulation of the problem. In the previous section, we describe the general form (13) of the dependence of σ_i on ε_{t-i} and σ_{t-i}. In this formula, the dependence on each input ε_{t-i} occurs through an expression $f_i(\varepsilon_{t-i})$ for some function $f_i(x)$ of one variable. Therefore, to describe the dependence on ε_{t-i}, we must describe the corresponding functions $f_i(x)$.

Scale invariance: a natural feature of econometric descriptions. The numerical value of each economic variable depends on the choice of a measuring unit. If we choose another measuring unit, the situation remains the same, but the numerical values change.

For example, for a US investor, it is natural to describe all the prices in dollars. For a European investor, it is equally natural to translate all the prices into Euros. If we replace the original unit with a new unit which is λ times smaller, then all numerical values need to be multiplied by λ.

The models should not change if we simply change the units. When we replace the original values ε_{t-i} by new numerical values $\varepsilon_{t-i} = \lambda \cdot \varepsilon_{t-i}$ of the same quantity, then each corresponding term $f_i(\varepsilon_i)$ is replaced with a new term $f_i(\varepsilon'_i) = f_i(\lambda \cdot \varepsilon_i)$. Thus, the overall contribution of all these terms changes from the original value $I = \sum\limits_{i=1}^{q} \alpha_i \cdot f_i(\varepsilon_{t-i})$ to the new value $I' = \sum\limits_{i=1}^{q} \alpha_i \cdot f_i(\lambda \cdot \varepsilon_{t-i})$.

It is reasonable to require that the relative quantity of different contributions does not change, i.e., that if two different sets $x_i^{(1)} \stackrel{\text{def}}{=} \varepsilon_{t-i}^{(1)}$ and $x_i^{(2)} \stackrel{\text{def}}{=} \varepsilon_{t-i}^{(2)}$ lead to the same contributions $I^{(1)} = I^{(2)}$, then after re-scaling, they should also lead to the same contributions I'. Thus, we arrive at the following condition:

Scale invariance: precise formulation of the requirement. Let the values $\alpha_1, \ldots, \alpha_q$ be fixed. Then, the functions $f_1(x), \ldots, f_p(x)$ should satisfy the following condition: if for two sets $x_1^{(1)}, \ldots, x_q^{(1)}$ and $x_1^{(2)}, \ldots, x_q^{(2)}$, we have

$$\sum_{i=1}^{q} \alpha_i \cdot f_i(x_i^{(1)}) = \sum_{i=1}^{q} \alpha_i \cdot f_i(x_i^{(2)}), \tag{14}$$

then for every $\lambda > 0$, we must have

$$\sum_{i=1}^{q} \alpha_i \cdot f_i(\lambda \cdot x_i^{(1)}) = \sum_{i=1}^{q} \alpha_i \cdot f_i(\lambda \cdot x_i^{(2)}). \tag{15}$$

Analysis of the problem: from scale invariance to the functional equation. For simplicity, let us start with the case when the values $x_i^{(2)}$ are very close to $x_i^{(1)}$, i.e., when $x_i^{(2)} = x_i^{(1)} + k_i \cdot h$ for some constants k_i and for a very small real number h. For small h, we have

$$f_i(x_i^{(1)} + k_i \cdot h) = f_i(x_i^{(1)}) + f_i'(x_i^{(1)}) \cdot k_i \cdot h + O(h^2). \tag{16}$$

Substituting the expression (16) into the formula (14), we conclude that

$$\sum_{i=1}^{q} \alpha_i \cdot f_i'(x_i^{(1)}) \cdot k_i \cdot h + O(h^2) = 0. \tag{17}$$

Dividing both sides by h, we get

$$\sum_{i=1}^{q} \alpha_i \cdot f_i'(x_i^{(1)}) \cdot k_i + O(h) = 0. \tag{18}$$

Similarly, the condition (15) leads to

$$\sum_{i=1}^{q} \alpha_i \cdot f_i'(\lambda \cdot x_i^{(1)}) \cdot k_i + O(h) = 0. \tag{19}$$

In general, the condition (18) lead to (19). In the limit $h \to 0$, we therefore conclude that for every vector $k = (k_1, \ldots, k_q)$, if

$$\sum_{i=1}^{q} k_i \cdot (\alpha_i \cdot f_i'(x_i^{(1)})) = 0, \tag{20}$$

then

$$\sum_{i=1}^{q} k_i \cdot (\alpha_i \cdot f_i'(\lambda \cdot x_i^{(1)})) = 0. \tag{21}$$

Functional equation: geometric analysis. The sum (20) is a scalar (dot) product between the vector k and the vector a with components $\alpha_i \cdot f_i'(x_i^{(1)})$. Similarly, the sum (21) is a scalar (dot) product between the vector k and the vector b with components $\alpha_i \cdot f_i'(\lambda \cdot x_i^{(1)})$. Thus, the above implication means that the vector b is orthogonal to every vector k which is orthogonal to a, i.e., to all vectors k from the hyperplane consisting of all the vectors orthogonal to a.

It is easy to see geometrically that the only vectors which are orthogonal to the hyperplane are vectors collinear with a. Thus, we conclude that $b = \delta \cdot a$ for some constant δ, i.e., that

$$\alpha_i \cdot f_i'(\lambda \cdot x_i^{(1)}) = \delta \cdot \alpha_i \cdot f_i'(x_i^{(1)}). \tag{22}$$

Dividing both sides by α_i, we conclude that

$$f_i'(\lambda \cdot x_i^{(1)}) = \delta \cdot f_i'(x_i^{(1)}). \tag{23}$$

Analysis of dependence and the resulting new differential equation. In principle, δ depends on λ and on values $x_i^{(1)}$. From the equation (23) corresponding to $i = 1$, we see that

$$\delta = \frac{f_1'(\lambda \cdot x_1^{(1)})}{f_i'(x_1^{(1)})}. \tag{24}$$

Thus, δ only depends on $x_1^{(1)}$ and does not depend on any other value $x_i^{(1)}$. Similarly, by considering the case $i = 2$, we conclude that δ can depend only on $x_2^{(1)}$ and thus, does not depend on $x_1^{(1)}$ either. Thus, δ only depends on λ, i.e., the condition (23) takes the form

$$f_i'(\lambda \cdot x_i^{(1)}) = \delta(\lambda) \cdot f_i'(x_i^{(1)}). \tag{25}$$

From the solution to the new differential equation to the solution of our original problem. It is known that every continuous function $f_i'(x)$ satisfying the equation (23) has the following form:

- $f_i'(x) = C_i^+ \cdot x^{a_i}$ for $x > 0$, and
- $f_i'(x) = C_i^- \cdot |x|^{a_i}$ for $x < 0$,

for some values $C_i^\pm$ and a_i; see, e.g., [1], Section 3.1.1, or [9]. (This result was first proven in [11].) For differentiable functions, the easiest way to prove this result is to differentiate both sides of (23) by λ, set $\lambda = 1$, and solve the resulting differential equation.

For the corresponding functions, the condition (25) is satisfied with $\delta(\lambda) = \lambda^{a_i}$. Since the function $\delta(\lambda)$ is the same for all i, the value a_i is therefore also the same for all i: $a_1 = \ldots = a_q$. Let us denote the joint value of all these a_i by a.

Thus, all the derivatives have $f_i'(x)$ are proportional to x^a. Hence, the original functions are proportional

- either to x^{a+1} (for $a \neq -1$)
- or to $\log(x)$ (when $a = -1$).

The additive integration constant can be absorbed into the additive constant α_0, and the multiplicative constants can be absorbed into a factor α_i.

Thus, without losing generality, we can conclude that in the scale invariant case, either $f_i(x) = x_i^a \cdot (1 + b \cdot I(x))$, or $f_i(x) = \log(|x|)$.

Conclusion. We have proven that the natural scale-invariance condition implies that each function $f_i(x)$ has either the form $\log(x)$, or the form $f_i(x) = x_i^a \cdot (1 + b \cdot I(x))$. This conclusion covers all the functions which are efficiently used to describe asymmetric heteroskedasticity:

- the function $f_i(x) = 1 + (\gamma_i/\alpha_i) \cdot I(x)) \cdot x^2$ used in Glosten's model;
- the function $f_i(x) = x^+ + (\alpha^-/\alpha^+) \cdot x^-$ used in Zakoian's model; and
- the function $f_i(x) = (1 + (\gamma_i/\alpha_i) \cdot \text{sign}(x)) \cdot |x|$ used in Nelson's model.

It is worth mentioning that this result also covers the functions $g_i(x) = x^2$ and $g_i(x) = \log(x^2) = 2\log(x)$ used to describe the dependence on σ_{t-i}.

Thus, the exact form of the dependence on ε_{t-i} has indeed been justified by the natural scale invariance requirement – as well as the dependence on σ_{t-i}.

Comment. It is worth mentioning that scale-invariance of the econometric formulas describing heteroskedasticity was noticed and actively used in [3]. However, our approaches are somewhat different:

- In [3], the econometric *formulas* were taken as *given*, and scale invariance was used to analyze heteroskedasticity tests.
- In contrast, we use scale invariance to *derive* the econometric *formulas*.

Acknowledgments.

This work was supported in part by NSF grant HRD-0734825 and by Grant 1 T36 GM078000-01 from the National Institutes of Health.

The authors are thankful to all the participants of the 2nd International Conference on Econometrics (Chiang Mai, Thailand, January 2009), especially to Hung T. Nguyen, for valuable discussions.

References

[1] J. Aczel, *Lectures on Functional Equations and Their Applications*, Dover Publ., New York, 2006.

[2] T. Bollerslev, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 1986, Vol. 31, pp. 307–327.

[3] J.-M. Dufour, L. Khalaf, J.-T. Bernard, and I. Genest, Simulation-based finite-sample tests for heteroskedasticity and ARCH effects, Journal of Econometrics, 2004, Vol. 122, No. 2, pp. 317–347.

[4] W. Enders, *Applied Econometrics Time Series*, John-Wiley & Sons, 2003.

[5] R. F. Engle, Autoregressive conditional heteroscedasticity with estimates of variance of united kingdom inflation, *Econometrica*, 1982, Vol. 50, pp. 987–1008.

[6] L. R. Glosten, R. Jagannathan, and D. E. Runkle, David E., On the relation between the expected value and the volatility of the nominal excess returns on stocks, *Journal of Finance*, 1993, Vol. 48, No. 5, pp. 1779–1801.

[7] D. Gujarati and D. Porter, *Basic Econometrics*, McGraw-Hill, 2008.

[8] D. B. Nelson, Conditional heteroskedasticity in asset returns: a new approach, *Econometrica*, 1991, Vol. 59, pp. 347–370.

[9] H. T. Nguyen and V. Kreinovich, *Applications of continuous mathematics to computer science*, Kluwer, Dordrecht, 1997.

[10] H. T. Nguyen and V. Kreinovich, Kolmogorov's Theorem and its impact on soft computing, In: R. R. Yager and J. Kacprzyk, *The Ordered Weighted Averaging Operators: Theory and Applications*, Kluwer, Boston, Massachusetts, 1997, pp. 3–17.

[11] J. Pexider, Notiz uber Funktionaltheoreme, Monatsch. Math. Phys., 1903, Vol. 14, pp. 293–301.

[12] J.-M. Zakoian, Threshold heteroskedastic models, *Journal of Economic Dynamics and Control*, 1994, Vol. 18, No. 5, pp. 931–955.