Pricing in the real estate market as a stochastic limit. Log Normal approximation

Oleg V. Rusakov, Michael B. Laskin, and Olga I. Jaksumbaeva

Abstract**—**We construct a stochastic model of real estate pricing. The method of the pricing construction is based on a sequential comparison of the supply prices. We proof that under standard assumptions imposed upon the comparison coefficients there exists an unique non-degenerated limit in distribution and this limit has the Log Normal law of distribution. The accordance of empirical distributions of prices to the theoretically obtained Log Normal distribution we verify by numerous statistical data of real estate prices from Saint-Petersburg (Russia). For establishing this accordance we essentially apply the efficient and sensitive test of fit of Kolmogorov-Smirnov. Basing on the world admitted standard of estimation prices in real estate market, we conclude that the most probable price, i.e. mode of distribution, is correctly and uniquely determined under the Log Normal approximation. Since the mean value of Log Normal distribution exceeds the mode – most probable value, it follows that the prices valued by the mathematical expectation are systematically overstated.

Keywords**—**Real estate market value; stochastic model of pricing; geometric Brownian motion; Sharpe parameter; applications of the Kolmogorov-Smirnov test of fit.

I. Introduction

"FOR the aim of establishing of the market value one determines the most probable price with the following necessary conditions. At this price the real estate object could be sold on the date of estimation in the open competitive market, when the parties to the deal behave reasonably, having full access to all necessary information, and the price of the deal is not affected by any force majeure circumstances". This is the accepted standard of estimation in real estate markets in many countries, particularly in Russian Federation [1].

Analogous formulations, based on a term of "the most probable price", are included in many others classic estimation standards, such as IVS (p.31 a) [2], TEGOVA (EVS p. 5.3.1) [3], USPAP (Standard rule 6-2, p.c.)[4], RICS (p.3.2.1) [5].

Mathematically, this means that the evaluation of the market price, stipulated by standards, essentially relies on a probabilistic nature of prices: price is a random variable. At the same time, the proper market value is such numerical characteristics of a distribution of the random variable (the price), which is called "mode" and it indicates the most probable value of the random variable.

In our paper, under sufficiently simple and natural assumptions, we establish the theoretical fact of weak convergence[1] of a process of sequential price comparisons to the Log Normal distribution. For establishing this fact we construct a stochastic model, where the geometric Brownian motion (gBm) is a limit (in a sense of convergence of finite dimensional distributions). We have exhaustive statistical data, which verify the mentioned fact. A number of key examples are examined. We establish a number of special properties for market values, which follow from the property of Log Normal distribution of prices. We use examples based on statistical data for the market prices of the existing residential real estate properties in the city of Saint-Petersburg as of 31.03.2014.

T. Ohnishi, T. Mizuno, C. Shimizu, T. Watanabe, Y. Saita [6]-[8] analyze data of building prices in Tokyo metropolitan area for a period from 1986 year. Their statistical analysis of cumulative distribution functions of the prices (which are grouped by the building size rule – size adjusted prices) show a defined agreement with the Log Normal Low. Let us cite them verbatim: "In sum, the presented results indicate that size adjusted prices follow a lognormal distribution for the quiet periods without extremely large fluctuations in prices". However, these authors do not consider *any stochastic pricing model* which leads to a Log Normal distribution. In paper [7] T. Ohnishi, T. Mizuno, C. Shimizu, T. Watanabe apply (briefly) the Central Limit Theorem in the Lindeberg's form to construct the Log Normal distribution pricing model in the real estate market.

P.Ciurlia, A.Gheno [9] solve a problem of estimation of option prices when the underlying asset is a real estate object. They directly postulate that the underlying asset prices follow the geometric Brownian motion. Thus, they automatically imply that real estate prices fit a Log Normal distribution.

This work was reported in the Second International Conference on Mathematics and Computers in Sciences and in Industry (MCSI), Malta, August 17, 2015.

O.V. Rusakov is with Mathematics and Mechanics Department, Saint-Petersburg State University, Saint Petersburg State University, St.Petersburg State University, 7/9 Universitetskaya nab., St. Petersburg, 199034 Russia; e-mail: o.rusakov@spbu.ru ; OViRusakov@yahoo.co.uk

M.B. Laskin is with Faculty of Economics, Saint-Petersburg State University, St. Petersburg, ul.Chaykovskogo, 62, 191194 Russia; e-mail: laskinmb@yahoo.com .

O.I. Jaksumbaeva is with Faculty of Economics, Saint-Petersburg State University, St. Petersburg, ul.Chaykovskogo, 62, 191194 Russia; e-mail: olgadzva@gmail.com .

[1] In the one-dimensional case: the weak convergence is equal to the convergence in distribution. However, in the given paper at times we exceed frames of the one-dimensional case, for instance, — in the key Lemma 1.

II. Assumptions for the Stochastic Model of GBM

We construct a stochastic pricing model, for the residential real estate market, based on the sequential comparison of objects of real estate. The coefficients of comparison are multiplicative, small, random, and correspond to one simple distribution (they are elementary and homogenous one, in certain sense). We show that for a sufficiently large number of comparisons[2] N, under standard assumptions on smallness and homogeneity of the coefficients of comparison, the Log Normal distribution of price[3] will form with high level of exactness. We suppose that the following simple and standard assumptions hold:

A. The coefficients of comparison are homogeneous ones, they are multiplicative and indicate by how much time the price of the next object of comparison is greater or lesser than the antecedent. The coefficients of comparison are homogeneous ones, they are multiplicative and indicate by how much time the price of the next object of comparison is greater or lesser than the antecedent.

B. The coefficients of comparison are totally independent: increases/decreases in price change independently of the previous steps of comparison.

Let us introduce a non-random value $S_0 \geq 0$, which we assume as the initial price of the object of comparison (the starting price). Let us consider a Log Normal distributed random variable X, with $E(\ln X) = \mu$ and $D(\ln X) = \sigma^2$, where μ is a real constant and $\sigma > 0$ is a strictly positive constant. Let us rewrite the random variable X in the following form

$$X = S_0 \exp\{(\mu - \ln S_0) + \sigma\, \mathrm{N}(0,1)\}, \qquad (1)$$

here and below $\mathrm{N}(0,1)$ denotes a standard normal random variable as well as the standard normal law of distribution.

Let us introduce the parameters of "ratio of "signal/noise", which indicate a relative risk. In fact, this is adopted modification of the Sharpe parameter [10]. We will interpret the value $\mu - \ln S_0$ as *a useful signal*,

$$R = R(\mu, S_0, \sigma) \overset{\Delta}{=} \frac{\mu - \ln S_0}{\sigma},$$

$$R_2 = R_2(\mu, S_0, \sigma^2) \overset{\Delta}{=} \frac{\mu - \ln S_0}{\sigma^2}. \qquad (2)$$

The process of sequential comparison is executed according to the following logic. We select some initial object of comparison carrying the price S_0. The selection has to be random.

Otherwise, S_0 should not be from the extreme values of the range of prices of available comparison objects.

The transition to the next object of comparison should be such that increase or decrease of the price is small. We consider the binary scheme, which is generated by an underlying sequence of Bernoulli trials. The price of the next object of comparison is determined by the following formula: $S_1 = S_0 Y_1$. Where Y_1 is a random variable, which takes two values: $U > 1$ (*Up*) and $0 < D < 1$ (*Down*) with the corresponding probabilities $0 < p < 1$ and $q = 1 - p$. At the same time, we assume that U and D takes values close to 1 (upwards or downwards, respectively). The transition to the next object takes place according to a recurrence rule, which provides that at each time the price is multiplied by an independent coefficient and that all the coefficients are identically distributed.

To prove that the price in the given recursive algorithm approaches a Log Normal distribution, we prove the normality of the limit of sums of logarithms of independent, equally distributed comparison coefficients Y. Due to the continuity and boundedness of exponents on finite intervals, this is enough to establish the necessary fact of log-normality of the limit law for distribution of prices (because convergence in distribution imply convergence of continuous and bounded functionals [11]).

It is important to note that probabilities p and q depend on the selection of the starting price of comparison S_0. We initially assume that there is a range (interval) of prices of compared objects, which depends on the quantity of objects N (involved in the process of sequential comparison). This quantity must be comparable with a size of the population (existing or assumed), and the objects involved in the process of sequential comparison must be distributed homogeneously among the population.

When the price S_0 is close to the minimum of price range the probability p to select the next object more expensive (i.e. probability p of the event $\{S_1 > S_0\}$) is greater than the probability q to take the next object more cheap (the event $\{S_1 < S_0\}$). At the same time, the lower (closer to the left bound of the range) the price S_0 is chosen, probability p more exceeds the probability q.

This way we assume dependence $p = p(S_0)$, $q = q(S_0)$. Let us also note that when the selection of starting comparison price S_0 is random, it becomes dependent on the number of objects N and consequently we have dependence of probabilities p and q on N. We assume the simple "logarithmically symmetrical" structure of the coefficients of comparison, i.e. $D = 1/U$.

For the further constructs, we use the following key stochastic lemma of "universal type".

Let random variables (η_j), $j \in \mathrm{N}$, be independent, identically distributed with zero mean and have a distribution dependent on a give natural n (arbitrary, but fixed). We normalize the random variables η standardly for CLT (Central

[2] Starting from 500, or around, for the approximation error about the level 5%, since such level of error is inversely proportional to square root of *N*.

[3] We consider price for 1 square meter.

Limit Theorem): $\xi_j = \left(1/\sqrt{n}\right)\eta_j$, $j \in \mathrm{N}$. We introduce the random variables ξ and η for denotation of the law of distribution of ξ_1 and η_1, respectively.

Further denotes and interpretation is as follows: the random variables as $\left(\xi_j\right)$ represent corresponding centered (i.e. after the subtraction of relevant expectations) logarithms of coefficients of comparison of the type $\ln Y$ and the number of these coefficients of comparison is n. The constant C_n, introduced in formula (3) below is the sum of expectations of logarithms (accumulated expectation of logarithms) of coefficients of comparison of the type $\ln Y$.

By the sequence of $\xi_1, \xi_2, \ldots, \xi_n$ we construct a continuous random broken line of the following type:

$$g_n(t) \overset{\Delta}{=} \sum_{j=1}^{[nt/T]} \xi_j + \left(\frac{nt}{T} - \left[\frac{nt}{T}\right]\right)\xi_{[nt/T]+1} + tC_n \tag{3}$$

where $t \in [0,T]$, terminal $T < \infty$, and C_n is a constant, dependent only on n. It is obvious that there exists a natural one-to-one map between the broken lines $\left(g_n\right)$ and the sequences $\left(\xi_1, \xi_2, \ldots, \xi_n\right)$.

Lemma 1. Let $\mathrm{E}\eta^2 < \infty$. Let random variables $\xi_j = \xi_j(n)$, $j = 1, \ldots, n$, and the constants C_n, σ_n satisfy the following conditions, as $n \to \infty$,

(i) $\quad \mathrm{E}\xi_1(n) = 0;$

(ii) $\quad \mathrm{E}\xi_1^2(n) = \dfrac{\sigma_n^2}{n} = \dfrac{\sigma^2}{n} T\left(1 + \overset{=}{o}(1)\right);$

(iii) $\quad C_n = C(1 + \overset{=}{o}(1)),$

where real C and $\sigma > 0$ are constants independent of n. Then the finite-dimensional distributions of random broken lines $\left(g_n\right)$ weakly converge, as $n \to \infty$, to finite-dimensional distributions of the random process $\sigma W_t + Ct$, $t \in [0,T]$, where W_t is the standard Brownian motion.

Detailed proof of Lemma 1 see in section 4.

Corollary 1 of Lemma 1. The sequence $\left(\exp\{g_n(T)\}\right)$ converges in distribution, as $n \to \infty$, to the Log Normal random variable $\exp\{C + \sigma\,\mathrm{N}(0,1)\}$.

Remark 1 to Lemma 1. Parameter $t \in [0,T]$ in the formulation of Lemma 1 one naturally interprets as a continuous time of execution of the process of comparisons. Involving more and more objects in the process of comparisons, fulfilling the conditions of homogeneity, we obtain (in limit) the process of Brownian motion for the logarithm of prices with the following main parameter: process of linearly accumulating volatility $\sigma^2 t = D\{\sigma W_t\}$, $t \in [0,T]$. In the case of non-homogeneity, dependence the volatility on time is $\sigma^2 = \sigma^2(t)$ (but it must be bounded in the mean-square sense) and dependence of the mean on time is $C = C(t)$. Thus we obtain the convergence of logarithms of price to the process

$$\int_0^t C(\upsilon)d\upsilon + \int_0^t \sigma^2(u)dW_u \,,\; 0 \le t \le T \,, \tag{4}$$

where dW_u – is a standard Gaussian "white noise". In this case the process of accumulated volatility is the integrated volatility as a function on upper limit t of integrals in (4).

Remark 2 to Lemma 1. Let for every natural n and for some $\delta > 0$ the moment $\mathrm{E}|\eta|^{2+\delta} < \infty$. Suppose that

(iv) $\quad \mathrm{E}\left|\xi_1(n)\right|^r = \underset{=}{O}(n^{-r/2})\,, \qquad r > 1.$

In clause *(iv)* it is not necessary to suppose the existence of all moments. It is to be understood that for every $r > 1$, such that $E\left|\xi_1(n)\right|^r < \infty$, condition *(iv)* is satisfied. Naturally, all $r \le 2 + \delta$ will satisfy *(iv)*. Under assumptions *(i)*-*(iii)* of Lemma 1 and assumption *(iv)* the continuous random broken lines $\left(g_n\right)$ weakly converge, as $n \to \infty$, in the functional space of continuous functions $C_{[0,T]}$, to the random process $\sigma W_t + Ct$, $t \in [0,T]$, where W_t is the standard Brownian motion. Thus, under assumptions *(i)*-*(iv)* functional limit theorem is fulfilled.

III. A PROCESS OF SEQUENTIAL COMPARISONS AND THE CORRESPONDING STOCHASTIC MODEL FOR PRICE ESTIMATION

Let us write down the structure of comparison step by step:

$$S_{j+1} = S_j Y_{j+1}, \quad j \le N \,, \tag{5}$$

where $\left(Y_j\right)$, $j = 1, \ldots, N$, – the sequence of i.i.d. random variables with the following distribution (given by the random variable Y)

$$Y \overset{\Delta}{=} \begin{cases} U, \; p \\ D, \; q \end{cases} = \begin{cases} U, & p(S_0) \\ \dfrac{1}{U}, & q(S_0) \end{cases} \tag{6}$$

Suppose that $U = u^{1/\sqrt{N}}$ for any u which is a slowly varying variable with respect to N (u is a constant, for instance).

Put

$$\xi_j^{'} = \ln Y_j \,, \qquad j \le N \,. \tag{7}$$

then

$$\xi_j = \xi_j^{'} - \mathrm{E}\xi_j^{'} = \ln Y_j - \mathrm{E}\ln Y_j \,, \qquad j \le N \,. \tag{8}$$

Put (in the notations for Lemma 1) $n = N$.

Let us suppose that the first and second central moments are as follows ($N \to \infty$)

$$\begin{cases} \mathrm{E}\xi_j^{'} = (\mu - \ln S_0)\dfrac{1}{N}\left(1 + \overset{=}{o}(1)\right); \\ D\xi_j^{'} = D\xi_j = \dfrac{\sigma^2}{N}\left(1 + \overset{=}{o}(1)\right); \qquad j \le N. \end{cases} \tag{9}$$

Here, the conditions of Lemma 1 are satisfied. In formulation of Lemma 1 we take $t = 1$, denote $\mathrm{E}\xi_j^{'} = C_j = C_j(N)$, and conclude that logarithms of prices (ξ_j) converges (weakly) to a normal distribution, as $N \to \infty$,

$$\ln\left\{S_0 \prod_{j=1}^{N} Y_j\right\} = \ln S_0 + \sum_{j=1}^{N} \xi_j + \sum_{j=1}^{N} C_j \Rightarrow \sigma\,\mathrm{N}(0,1) + \mu\,. \tag{10}$$

Remark 1. Let us note that in the limiting expression (10) the distribution of the comparison coefficient of type Y does not necessarily belong to a binary distribution law. The only requirement is that the conditions of Lemma 1, related to independence and moments, are satisfied. Therefore, we can talk about the invariance of the method of proof as regards to the selection of law of distribution of comparison coefficients.

Remark 2. Due to Lemma 1, accumulated sums converge to the Brownian motion and we obtain a Log Normal distribution in the case when we have sub-sample or increased sample volume as $[Nt]$, $0 < t < T < \infty$, $[\cdot]$ denotes integer part. The only important thing here is that the coefficients of comparison must be independent and homogeneously small in certain terms. In such cases, we only have the change in variance of normal distribution of the logarithm in accordance with the properties of the standard Brownian motion, $DW_t = t$, $t \ge 0$. In the case of non-homogeneity, we have a corresponding diffusion process with an accumulated volatility (See Remark to Lemma 1).

Remark 3. It is well known that convergence in distribution does not necessitate the convergence of moments (expectation, variance) and some other properties of the distribution. However, such moment characteristics will converge (under condition of weak convergence of distribution) in case, when they are uniformly bounded. This condition is automatically fulfilled for our model, since the logarithms of all prices are bounded by a single constant, which can be very large.

Let us consider the following problem of invariance of probabilities in distribution (6) of the coefficients (Y_j) in (5), with respect to selection of the primary object with the initial price S_0 in our process of comparison. We apply known the elementary formulae for the variance and expectation of binary random variables and calculate the moments of the random variables of type $\xi_j^{'}$

$$\mathrm{E}\xi_j^{'} = p\ln U + d\ln D = (2p-1)\ln U, \qquad j \le N, \tag{11}$$

$$D\xi_j^{'} = D\xi_j = (\ln U - \ln D)^2 pq = 4p(1-p)(\ln U)^2 \tag{12}$$

Setting equal the expectation (11) and variance (12) to the corresponding expressions in (9) we derive, neglecting members of smaller level (as $N \to \infty$), the following system of two equations (with respect to unknown p and S_0)

$$\left.\begin{aligned} (2p-1)\ln U &= (\mu - \ln S_0)\frac{1}{N}; \\ \sqrt{p(1-p)}\,2\ln U &= \frac{\sigma}{\sqrt{N}}. \end{aligned}\right\} \tag{13}$$

We divide in (13) the first equation by the second equation and obtain the following Proposition 1, which relates probability p to the coefficient of relative risk R, determined by (2), (and accordingly: p to S_0 in the case of known values of the parameters μ and σ^2)

Proposition 1. The probability of increase in price p in the process of sequential comparisons depends only on the coefficient of relative risk and the quantity of objects of comparison:

$$\frac{1}{2}\frac{p - (1-p)}{\sqrt{p(1-p)}} = \frac{1}{\sqrt{N}}\frac{\mu - \ln S_0}{\sigma} = \frac{R}{\sqrt{N}} \tag{14}$$

Comments to Proposition 1. Equation (14) has a term, decreasing to zero with the rate of square root from N. This is a result of application of the CLT to the sequence of Bernoulli trials: deviation of number of successes from the expected number of successes falls within the interval proportional to the square root from N.

Let us denote the left-hand side in (14) as a function $l(p)$, where $p \in (0,1)$. Consider behavior of this function

$$l(p) \overset{\Delta}{=} \frac{1}{2}\left(\sqrt{\frac{p}{1-p}} - \sqrt{\frac{1-p}{p}}\right) \tag{15}$$

Note that the function $l(p)$ is continuous (moreover, infinitely-differentiable), with strictly monotone growth, tending to minus infinity as $p \to 0$, equals zero when $p = 1/2$ (i.e. in the case of the symmetrical random walk) and tends to plus infinity as $p \to 1$. This shows that in equation (15) we have a unique solution $p \in (0,1)$ for the case of any real left part.

Moreover, when the values of parameters μ and σ^2 and the sample volume N are known (or given), we can calculate $p = p(S_0)$ for any positive value of S_0, – and conversely. Since we are working with a sample (of real estate prices), we have the volume N of the sample, therefore we can always compute sample mean and sample variance, put them into (14) instead of μ and σ^2, respectively and obtain a straight (calculable) ratio between S_0 and $p = p(S_0)$ due to (15).

Note, that D.Bilkova ([12], [13]) solves a problem of estimation of parameters (θ, μ, σ) for the distribution X (defined by (1)) with a shift $\theta > 0$,

$$X_\theta = \theta + S_0 \exp\{(\mu - \ln S_0) + \sigma\,\mathrm{N}(0,1)\}, \tag{16}$$

Author of these papers suppose that “Three-parametric Log Normal distribution was a fundamental theoretical distribution for estimations of individual data on net annual household income”. So, it may follow from the fact of Log Normality of prices of the corresponding houses.

We also remark the average price for gBm over the interval $[0,T]$ has a quite large correlation R with respect to the price S_T. It follows from the results of B.J.C.Baxter and R.Brummelhuis [14]:

“The correlation coefficient R is typically close to unity: typical values of μ, σ, and T produce values of R in the 0.8-0.9 range.”

This results of B.J.C.Baxter and R.Brummelhuis allows us to conclude, that our model has (in some weak sense) an ergodic property which connect the statistical average value (over the ensemble of objects of comparisons) with the price of “a typical object”.

IV. Detailed Proof of Lemma 1

Let $f(x)$ a characteristic function of distribution of the random variable ξ_1. First we note that under conditions *(i), (ii)* for any fixed real x the following standard decomposition (as for CLT) is valid

$$\ln f(x) = -\frac{1}{2}\frac{\sigma^2 T}{n}x^2 + \overset{=}{o}\left(\frac{1}{n}\right), \qquad n\to\infty . \tag{17}$$

Let us consider an arbitrary (not necessary uniform) partition of interval $[0,T]$ which is given by the points $0=t_0<t_1<\cdots t_m<T$, $m\in \mathrm{N}$. We calculate distribution of the random vector with the components equal the values at these points of the random broken line g_n

$$\varphi=\varphi_n(m)\overset{\Delta}{=}\left(g_n(t_1),\ldots,g_n(t_m)\right), \tag{18}$$

where $(g_n(t))$, $n\in\mathrm{N}$, are defined by (3). For simplicity further we omit the index n in denotations of the random function $g_n(t)=g(t)$.

Let us denote t_j^*, $j=1,\ldots,m$, the nearest (from the left side to the point t_j) node of the uniform partition with the rank n of the interval $[0,T]$:

$$t_j^*\overset{\Delta}{=}\left[\frac{t_j n}{T}\right]\frac{T}{n}=t_j-\left\{\frac{t_j n}{T}\right\}\frac{T}{n}, \tag{19}$$

here $\{a\}$ denotes the fractional part of a number a. It is not difficult to derive from (3) the following equality

$$g(t_j)-g(t_j^*)=\left(\frac{n}{T}t_j-\left[\frac{n}{T}t_j\right]\right)\left(\xi_{[nt_j/T]+1}+C_n\frac{T}{n}\right). \tag{20}$$

Since the first term in the right part of this equality is uniformly (by n) bounded by 1, and the second term has the expectation and the variance tending to zero (as $n\to\infty$), so thanks the classical Cnebyshev inequality we easy establish the following convergence in probability

$$\left|g(t_j)-g(t_j^*)\right|\xrightarrow{P}0, \qquad n\to\infty, \tag{21}$$

Note, that convergence in distribution to a constant includes the corresponding convergence in probability.

So we conclude that if the vector φ has a limit distribution, then this distribution coincides with the limit distribution of the following random vector

$$\varphi^*=\varphi_n^*(m)\overset{\Delta}{=}\left(g(t_1^*),\ldots,g(t_m^*)\right)\overset{\Delta}{=}\left(g\left(\frac{k_1T}{n}\right),\ldots,g\left(\frac{k_mT}{n}\right)\right), \tag{22}$$

where $k_j\overset{\Delta}{=}[nt_j/T]$, $j=1,\ldots,m$.

Let us consider the random vector ψ with components which are equal to increments of the vector φ^*:

$$\psi=\psi(n)\overset{\Delta}{=}(\psi_1,\ldots,\psi_m) \tag{23}$$

with the corresponding components, $j=1,\ldots,m$,

$$\psi_j\overset{\Delta}{=}g\left(k_j\frac{T}{n}\right)-g\left(k_{j-1}\frac{T}{n}\right). \tag{24}$$

It is obvious that

$$\psi_j=\sum_{i=k_{j-1}+1}^{k_j}\xi_i+C_n\left(k_j-k_{j-1}\right)\frac{T}{n}. \tag{25}$$

Moreover, $(\psi_1,\ldots,\psi_m)$ are independent random variables, and even more: each component of the vector ψ consists of sums of independent random variables. Thus for any $j=1,\ldots,m$ the characteristic function f_j of the random variable ψ_j may be written in the following form

$$f_j(x;n)=f_j(x)=\exp\left\{i\,xC_n\left(k_j-k_{j-1}\right)\frac{T}{n}\right\}f^{\,k_j-k_{j-1}}(x), \tag{26}$$

where $f(x)$, $x\in\Re$, is a characteristic function of the random variable ξ_1.

Since the following limit is obvious for $j=1,\ldots,m$

$$\lim_{n\to\infty}\left(k_j-k_{j-1}\right)\frac{T}{n}=t_j-t_{j-1}, \tag{27}$$

then for every fixed $x\in\Re$ for the characteristic function $f_j(x)$ for every $j=1,\ldots,m$ the following limiting relation is valid

$$\lim_{n\to\infty}f_j(x;n)=\exp\{i\,xC_n(t_j-t_{j-1})\}\exp\left\{-\frac{\sigma^2}{2}(t_j-t_{j-1})x^2\right\}. \tag{28}$$

Here we apply the limit relation *(iii)* for the sequence (C_n) from assumptions of Lemma 1, and we also use the representation of logarithm of the characteristic function $f(x)$ of the random variable ξ_1.

Next step we apply a classical theorem on the continuous correspondence between characteristic functions and distributions, and we derive that distributions of the random vector ψ (indexed by n) weakly converge, as $n\to\infty$, to the m-dimensional normal distribution with the vector of the corresponding expectations

$$\left(t_1-t_0,t_2-t_1,\ldots,t_m-t_{m-1}\right)C, \tag{29}$$

and the matrix of covariances

$$\delta_{ij}\,\sigma^2\left(t_j - t_{j-1}\right); \qquad i, j = 1, \dots, m, \tag{30}$$

where $\delta_{ij} = 0$, as $i \neq j$, and $\delta_{ij} = 1$, as $i = j$, – denotes the Kronecker symbol.

Finally, we examine a proximity of the vectors φ and φ^*. It is easy to obtain that their distributions coincides up to an infinitely small term, as $n \to \infty$. At the same time, the vector φ^* one can construct from the vector ψ by summing up of independent components (of ψ), so it is not difficult to calculate the joint limit distribution of the vectors φ and φ^*. It is: *m-dimensional gaussian distribution with the vector of means $(t_1, \dots, t_m)C$ and the covariance matrix $\sigma^2 \min\left(t_i, t_j\right)$*, $i, j = 1, \dots, m$.

The obtained characteristics fully corresponds to the well known characteristics of finite dimensional distributions of the (gaussian) Wiener process with the scale parameter σ and the shift Ct. Therefore, we establish the following fact of convergence. Finite dimensional distributions of the process defined by the random broken lines $g_n(t)$, weakly converge, as $n \to \infty$, to the finite dimensional distributions of the process

$$\sigma W_t + C\,t, \qquad t \in [0, T], \tag{31}$$

where W_t – is a standard Wiener process (Bm). Lemma 1 is proved.

V. Numerical Results

The splitting of Saint-Petersburg city into administrative districts shows that the municipal districts have significantly homogenous residential real estate units. As a result, the distributions of price for one square meter in these districts one can represent by the Log Normal Distribution Law.

Applications of the introduced model we illustrate by the following examples for real estate prices on offer, dated 31.03.2014, for several regions of Saint-Petersburg city.

In Fig.1 below one can see the histogram and fitting curve of density of the Log Normal distribution for the Admiralty district of Saint-Petersburg. Corresponding statistic values: volume of the sample: $N = 403$; estimated parameters of Log Normal law: $\mu = 4.57$, $\sigma = 0.226$; arithmetic average over the sample: 99.799 thousand RUR per 1 m^2; estimated market value (mode of the Log Normal approximation): 91.983 thousand RUR per 1 m^2; p-value for the Kolmogorov-Smirnov test of fit: 0.3245.

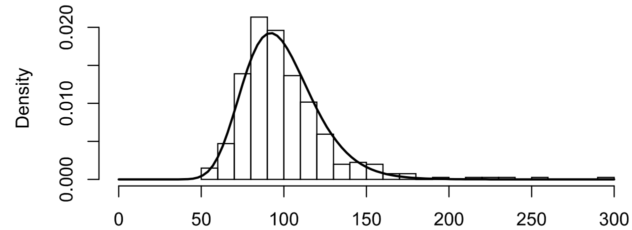

Fig.1. Admiralty district of Saint-Petersburg.

In Fig.2 below one can see the histogram and fitting curve of density of the Log Normal distribution for the Central district of Saint-Petersburg. Corresponding statistic values: volume of the sample: $N = 560$; estimated parameters of Log Normal law: $\mu = 4.921$, $\sigma = 0.34$; arithmetic average over the sample: 145.000 thousand RUR per 1 m^2; estimated market value (mode of the Log Normal approximation): 122.168 thousand RUR per 1 m^2; p-value for the Kolmogorov-Smirnov test of fit: 0.3038.

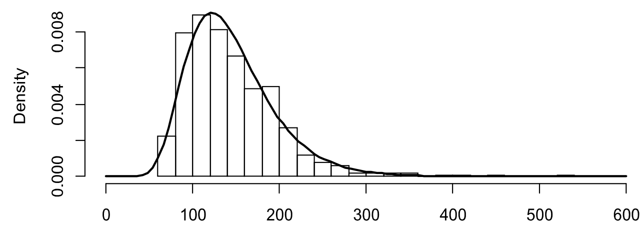

Fig.2. Central district of Saint-Petersburg.

In Fig.3 below one can see the histogram and fitting curve of density of the Log Normal distribution for the Vasileostrovsky district (Vasilevsky Island) of Saint-Petersburg. Corresponding statistic values: volume of the sample: $N = 697$; estimated parameters of Log Normal law: $\mu = 4.669$, $\sigma = 0.23$; arithmetic average over the sample: 109.448 thousand RUR per 1 m^2; estimated market value (mode of the Log Normal approximation): 101.099 thousand RUR per 1 m^2; p-value for the Kolmogorov-Smirnov test of fit: 0.110.

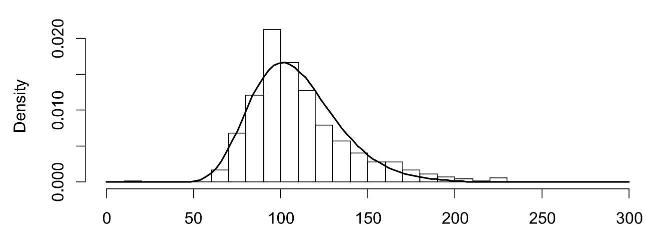

Fig.3. Vasileostrovsky district of Saint-Petersburg.

In Fig.4 below one can see the histogram and fitting curve of density of the Log Normal distribution for the Stalin period buildings in Saint-Petersburg. Corresponding statistic values: volume of the sample: $N = 647$; estimated parameters of Log Normal law: $\mu = 4.606$, $\sigma = 0.20$; arithmetic average over the sample: 102.105 thousand RUR per 1 m^2; estimated market value (mode of the Log Normal approximation): 96.159 thousand RUR per 1 m^2; p-value for the Kolmogorov-Smirnov test of fit: 0.342.

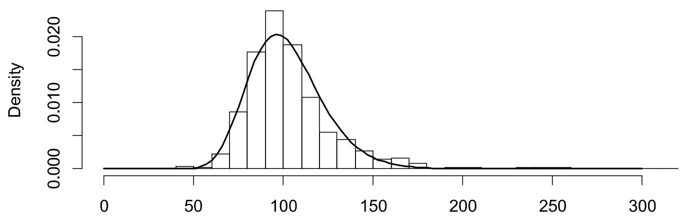

Fig.4. Stalin period buildings in Saint-Petersburg.

VI. Conclusion

A simple analysis of the constructed above stochastic model and the resulting formulae enable us to make the following important practical conclusions.

Conclusion 1. The mean of Log Normal law of distribution is $\exp\left(\mu+\sigma^2/2\right)$, and the mode is $\exp\left(\mu-\sigma^2\right)$ and it is easy to see that the ratio of mean to the mode is not dependent on μ, it is always greater than 1 and equals $\exp\left((3/2)\sigma^2\right)$. Expanding the exponent, at small values of σ we obtain that this value is asymptotically equal to $1+(3/2)\sigma^2$. It means that the usage of arithmetic averages for the evaluation of market value (most probable value – is mode!) leads to systematic error tending to overstating, expressed in the relative error: $\exp\left((3/2)\sigma^2\right)-1$ (×100%). For example when $\sigma=0.258$ the systematic error is about 10.5%.

Conclusion 2. Sometimes in surveys instead of arithmetic average (estimation of expectation) is used geometric average (estimation of median). Since the median of Log Normal law equals $\exp(\mu)$, therefore the ratio of median to mode does not depend on μ, is always greater than 1 and equals $\exp\left(\sigma^2\right)$. Expanding the exponent in series we obtain, under small σ, that the examined ratio is asymptotically equal to $1+\sigma^2$. It means that the usage of geometric averages to estimate market value leads us to systematic error tending to overstating, expressed in the relative error: $\exp\left(\sigma^2\right)-1$ (×100%). Particularly, when $\sigma=0.258$ the systematic error is about 6.9%.

Appendix

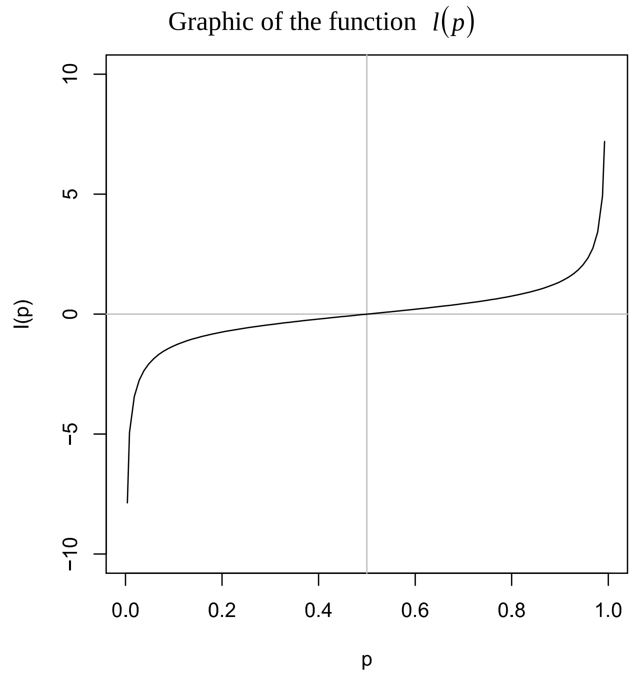

Fig.5. Graphic of the function *l*.

References

[1] Federal Standard of Survey (Valuation) № 2 "The purpose of valuation and types of prices". Decree of the Ministry of Economic Development of the Russian Federation №255 dated at 20.07.2007. Registered in the Ministry of Justice of RF 23.08.2007.

[2] International Valuation Standard Council, London, 2013. URL:http://www.valuersinstitute.com.au/docs/professional_practice/International%20Valuation%20Standards%202013.pdf

[3] European Valuation Standards, 7thedition, Brussels, 2012. URL: http://www.tegova.org/en/p4fe1fcee0b1db

[4] Uniform Standards of Professional Appraisal Practice, 2014–2015 edition, Annapolis, Maryland, USA, 2014.

[5] Royal Institution of Chartered Surveyors Valuation Professional Standard, London, 2014. URL:http://www.rics.org/uk/shop/RICS-Valuation-Professional-Standards-Red-Book-2013-19750.aspx

[6] T. Ohnishi, T. Mizuno, C. Shimizu, T. Watanabe, "On the evolution of the house price distribution," Columbia Business School, Center of Japanese Economy and Business, Working Paper Series, no. 296, 2011.

[7] T. Ohnishi, T. Mizuno, C. Shimizu, T. Watanabe, "Detecting Real Estate Bubbles: A New Approach Based on the Cross Sectional Dispersion of Property Prices," Paper presented at PKU-Fordham-JBF Conference on "Risk Management and Reform of Bank Regulation," April 22–23, 2013, Beijing. CARF Working Paper Series, CARF-F-313, April 2013.

[8] Y. Saita, C. Shimizu, T. Watanabe, "Aging and Real Estate Prices: Evidence from Japanese and US Regional Data," International Journal of Housing Markets and Analysis, Vol. 9 Issue 1, 2016, pp. 66–87.

[9] P. Ciurlia, A. Gheno "A model for pricing real estate derivatives with stochastic interest rates," Mathematical and Computer Modelling, vol.50, 2009, pp. 233–247.

[10] W. R. Sharpe, "The Sharpe Ratio," Journal of Portfolio Management 21(1), 1994, pp. 49–58.

[11] P. Billingsley, *Convergence of Probability Measures.* John Willey & Sons, New York, 1968.

[12] D. Bilkova, “Lognormal distribution and using l-moment method for estimating its parameters,” International Journal of Mathematical Models and Methods in Applied Sciences, 6(1), 2012, pp. 30–44.
[13] D. Bilkova, “Comparison of the accuracy of l-moments, tl-moments and maximum likelihood methods of parameter estimation,” International Journal of Mathematical Models and Methods in Applied Sciences, no.9, 2015, pp. 55–89.
[14] B.J.C.Baxter, R.Brummelhuis, “Functionals of exponential brownian motion and divided differences,” Journal of Computational and Applied Mathematics, 236(4), 2011, pp. 424–433.
[15] O.V. Rusakov, O.I. Jaksumbaeva, A.A. Ivakina, M.B. Laskin, “Pricing in the Real Estate Market as a Stochastic Limit. Log Normal Approximation,” *IEEE Conference Publications*, 2015 Second International Conference on Mathematics and Computers in Sciences and in Industry (MCSI), 2015, pp. 235–239.