SEMIH YON
Istanbul Technical University, Turkey, yon@itu.edu.tr

CAFER ERHAN BOZDAG
Istanbul Technical University, Turkey, bozdagc@itu.edu.tr

TEST OF LOG-NORMAL PROCESS WITH IMPORTANCE SAMPLING FOR OPTIONS PRICING

Abstract:

Log-normal process and martingale restriction bring some bias on the premium for option pricing models. It is possible to reduce the bias by adding more parameters like jump diffusion, stochastic volatility or regime switching. In this case closed form solutions and numerical approximations suffer from the dimension of the problem. Monte Carlo integration then appears to be unique solution for high dimensional calculations. However variance of the output of interest should be decreased in Monte Carlo applications in order to have confident results. The method of Importance Sampling can be used in an attempt to reduce variance. In this study we test the log-normal process for options pricing via Importance Sampling Monte Carlo. Our analysis is based on the theory of variance reduction and we don't have any empirical data. Numerical results indicate that the risk neutral density should be substituted in the range of moneyness.

Keywords:

Options pricing, lognormal process, variance reduction, importance sampling, moneyness

JEL Classification: G13, G00, C15

1. INTRODUCTION

In developed financial markets firms and individuals seek new methods to minimize the risk arises from their transactions. Options allow investors to control the risk level when included to the portfolios. Huge amount of transactions in the options market makes options pricing one of the most attractive topics. Very famous Black and Scholes (1973) model lightened options pricing process. Their assumptions have been analyzed enormously. In this study we test the log-normal process which is found to be restrictive in some cases.

It is impossible to specify the probability distribution of risky asset returns. Options pricing is based on a different probability space that calculations are done with respect to arbitrage free principle and risk neutral pricing. In this probability space asset price is simply an expectation of the discounted measure of its terminal value. The expectation is taken under an equivalent martingale measure which is a mapping of the original probability distribution of asset returns. Arbitrage free principle provides existence of an equivalent martingale measure which is not unique if the market is incomplete, Liu and Zhao (2013). It is assumed that logarithmic returns have an equivalent normal distribution in Black/Scholes model. Hence, distribution of the asset price becomes log-normal. Merton (1976) added price jumps to the log-normal process. However final model cannot prevent some bias on the premium which increases with the maturity of the option. There are two main reasons for the bias. First one is market crashes not reflected by log-normal process with constant volatility. Hull and White (1987) introduced basic solution to stochastic volatility models excluding correlation between the volatility and the price of the underlying asset. Heston (1993) and Aït-Sahalia and Kimmel (2006) tried to find closed form solutions for general stochastic volatility models. Second bias effect comes from the market frictions such as transition costs and bid-ask spread of assets. Longstaff (1995) analyzed empirical data and showed that bias is greater for out-the money and in-the money options. Longstaff (1995) calculated implied volatility of the S&P index call options for two years and found the result that implied volatility of the S&P index options has a smile pattern. This is known as volatility smile anomoly and studied by various authors like Rubinstein (1994), Neumann (1998), Stein (1989) and Jackwert and Rubinstein (1996). Although Black/Scholes model imposes to observe the current underlying price from the market, Longstaff (1995) relaxed the underlying S&P index values and had the observation of implied index values being higher than the actual index values in 99 percent of his data which simply shows that it is more expensive to purchase the underlying asset from options market than the stock market. This is obviously the case of more transaction costs of options market. But Jackwert and Rubinstein (1996) showed that even the transaction costs remain constant the volatility smile happens to have different patterns for options with different underlying assets which proves that the only reason for the bias is not market frictions. Longstaff (1995) claimed if it is allowed in calculations to have the underlying asset price 0.4 percent more than its market value greater precision is obtained in means of smaller bias for option prices. Longstaff (1995) defined this basic assumption as martingale restriction. Estimating the implied index value and the implied volatility is the same as estimating the first and the second moments of the risk neutral density which is assumed to be log-normal in most cases. Hence, diversification of the model results and empirical data is mostly caused by the log-normal process itself. Neumann (1998) used two log-normal distributions as a mixed distribution to fit empirical data better. Neumann (1998) calculated the parameters of the mixed log-normal distribution with the least squares error technique so as to minimize the diversification. Most of the articles in the area refer to the term selected for the input data. Then the analysis gets dependent to the market events. The contribution of our study is that we do not use empirical data to test the risk neutral density.

Other advances to come up with the bias of Black-Scholes model are based on regime switching models. Bastani et al, (2013) study on American options with a radial basis collacation method. Boyle and Draviam (2007) studied on exotic options under regime switching model. Liu and Zhao (2013) deal with lattice methods for two underlying assets in regime switching model. Single risk-neutral density is not enough to represent the dynamics of option prices. Therefore randomly changing combination of Lévy processes included to the models. Brownian motion is the only Lévy process having continuous patterns. On the other hand, regime switching models with general Lévy processes are discrete realizations of the actual process whose states are determined by a continuous time Markov chain. Thus, Markov chain model allows to specify long run equilibrium probabilities. The rate of return of the regime switching models consisting of a number of Lévy processes with different parameters converges to the expected risk neutral return.

Remainder part of the paper is organized as follows. Section 2 is devoted to the options pricing basis with log-normal process. Section 3 is devoted to the Monte Carlo integration framework and Importance Sampling (IS) technique. Section 4 is concerned with numerical study. Firstly key factors in our simulation model are defined. Numerical results are discussed next. Black/Scholes model, Monte Carlo and IS Monte Carlo results are compared in the sense of variance reduction capability. Then statistical test results are displayed. And finally conclusions are set.

2. OPTIONS PRICING IN CLOSED FORM

Black and Scholes (1973) model is based on the main assumption of normal distributed logarithmic returns. The underlying asset price follows a geometric Brownian motion which is also called log-normal process. Then underlying price dynamics were reflected by a partial differential equation (PDE) as

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{2.1}$$

where $S_t \in Q^+$ is the spot price, μ is annual drift, σ is annual volatility of the underlying and dW_t is Wiener process. One easy way to show the derivation of this PDE is as follows: For one period risky asset price can be expressed as

$$S_t = S_{t-1} u_t \tag{2.2}$$

where $u \in \Box^{+}$ is a random variable which includes all economic information to change the price. Next step is to have logarithm of both sides and to start from the initial state

$$\log S_t = \log S_0 + \sum_{k=1}^{t} \log u_k. \tag{2.3}$$

If we have $\log u = \xi$ in equation (2.3) as a normal random variable with $\xi \sim \mathrm{N}(\mu, \sigma^2)$ as imposed in Black and Scholes (1973) model, logarithmic price also becomes a normal random variable. The expectation and variance of the logarithmic price is $\mathrm{E}[\log S_t] = \log S_0 + \mu t$ and $\mathrm{Var}[\log S_t] = \sigma^2 t$, respectively. Then standard normal *z* can be expressed for stochastic random variable $X_t = \log\frac{S_t}{S_0}$ as

$$z = \frac{X_t - \mu t}{\sigma\sqrt{t}} \tag{2.4}$$

For stochastic random variable $X_t = \log S_t - \log S_0$ equation (2.4) can be arranged as

$$\log S_t - \log S_0 = \mu t + \sigma\sqrt{t} z_i \tag{2.5}$$

where subscript *i* refers to the standard normal random number here. If we have differential for both sides in equation (2.5) we get

$$d \log S_t = \mu dt + \sigma dW_t \tag{2.6}$$

since $\sqrt{t} z_i$ in equation (2.5) is a random variable satisfying Brownian motion $\big(B(t) - B(0) \sim \mathrm{N}(0, t)\big)$ and therefore can be substituted with dW_t. Finally in equation (2.6) differential of logarithmic price is substituted with $d \log S_t = \frac{dS_t}{S_t}$ and equation (2.1) is found.

Parabolic PDE has a boundary at expiration time $t = T$ which serves as the option price. Payoff function for call option is $max(S_t - K, 0)$ where *K* is exercise price. When $S_t < K$ call option pays off zero. And payoff function for put option is $max(K - S_t, 0)$. Call option and put option prices were calculated by solving PDE in Black and Scholes (1973).

3. OPTIONS PRICING WITH MONTE CARLO SIMULATION

Option price (so called premium) is the discounted value of the payoff under a riskless interest rate. Payoff is simply an expectation of the return determined by stochastic price vector in arbitrage free environment. E.g European put option premium is calculated with

$$V_p(S, T) = E[e^{-rT}(K - S_T)^+]. \tag{3.1}$$

The terminal value of the underlying price should be identified for the use of equation (3.1). In option pricing models future price of the underlying is based on a random walk. It is possible to generate underlying price vector with an equivalent martingale measure. The future price is calculated with the following formula for log-normal process.

$$S_{i+1} = S_i e^{\left(r - \frac{\sigma^2}{2}\right)\Delta t + \sigma\sqrt{\Delta t} z_i}, z_i \sim \mathrm{N}(0,1) \tag{3.2}$$

where r is the riskless interest rate. S_0 is observed from the market. Time frame Δt is set in years and might have $\Delta t = 1/252$ for working days per year in case of daily closing prices are simulated.

3.1 Monte Carlo Integration

Monte Carlo integration technique is widely used in derivatives pricing. Many problems can be formulated as integrals over a single model distribution or highly multi-modal distributions in result of expectations which can be shown as

$$\theta_f = E_f[q(\boldsymbol{x})] = \iiint_{R^d} q(\boldsymbol{x}) f(x) d\boldsymbol{x} \tag{3.3}$$

where $q(\boldsymbol{x})$ is a real valued function. The notation θ_f, E_f denotes that the expectation is taken with respect to density $f(\cdot)$ which belongs to the *d*-dimensional probabilistic state space Ω. If it is hard to find a closed form solution to equation (3.3) Monte Carlo simulations can be warranted to provide approximate results. Simulations driven by random inputs will produce random outputs. And those random outputs are the estimation of the exact results. The accuracy of this estimation strongly depends on quality of sampling which can be improved in two ways:

- increasing the cardinality of sampling or,
- introducing some kind of selection rules that make it more representative.

The first choice is commonly known limited way whereas the second requires to apply some special techniques. One of them is explained in the next section.

3.2 Importance Sampling as a Variance Reduction Technique

In Monte Carlo applications variance of the output random variable should be reduced without disturbing its expectation, which means smaller confidence intervals. Importance Sampling (IS) improves the quality of sampling and used for variance reduction purposes. IS introduces definite selection rules to generate most likely configurations to obtain more accurate values of statistical averages. Certain values of the input random variables in a simulation have more impact on the parameter being estimated than others. If these important values are emphasized by sampling more frequently, then the estimator variance could be reduced. Hence, the basic methodology is to choose a new distribution which encourages the important values. Yön and Goldsman (2006) deal with some useful biasing methods. This use of a biased distribution will result in a biased estimator. However, the simulation outputs are weighted to correct for the use of the biased distribution, and this ensures that the new IS estimator is unbiased, Broadie and Glasserman (1997). IS can be carried out as in the following way:

$$\theta_g = E_g\left[q(\boldsymbol{x})\frac{f(\boldsymbol{x})}{g(\boldsymbol{x})}\right] = \iiint_{R^d} q(\boldsymbol{x})\frac{f(\boldsymbol{x})}{g(\boldsymbol{x})} g(x) d\boldsymbol{x}, \tag{3.4}$$

Random samples are generated from $g(\cdot) \in \Omega$ which is called IS density. $g(x)$ enables to calculate the correction factor $\frac{f(x)}{g(x)}$. Correction factor is sometimes called weight function. Based on sample weights accumulated during sampling the correction factor compensates for statistical fluctuations and lead to a lower variance. In equation (3.4) the IS density $g(x)$ should assign higher probabilities to important region while holding $\theta_f = \theta_g$, Yön (2007). Then the estimator can be calculated as

$$\widehat{\theta_g} = \frac{1}{N}\sum_{i=1}^{N}\left(q(\boldsymbol{x_i})\prod_{j=1}^{d}\frac{f(x_j)}{g(x_j)}\right) \tag{3.5}$$

where *N* is the replication number and *d* is the dimension of the multivariate underlying distribution. Note that $f(\cdot)$ and $g(\cdot)$ are two independent densities. Finally Mean Squared Error (MSE) of the estimator is calculated in usual form

$$MSE = \frac{\sum_{i=1}^{N}\left(\theta_i - \widehat{\theta_g}\right)^2}{(N-1)} \tag{3.6}$$

The successful IS density leads lower possible MSE. Detailed features of IS densities were given at Yön (2007) and Broadie and Glasserman (1997).

4. NUMERICAL RESULTS

We test log-normal process from a variance reduction point of view by nominating Importance Sampling (IS) technique. We first carry out crude Monte Carlo simulation and then run IS for the same input variables. In order to measure the performance of the log-normal process we define two key factors. The first one is Variance Reduction Factor (VRF) which is the ratio of variance outputs of both trials. We calculate VRF by dividing the Mean Squared Error (MSE) of crude Monte Carlo simulation (using log-normal density) to the MSE of IS Monte Carlo simulation (using IS density)

$$VRF = \frac{MSE_{cMC}}{MSE_{ISMC}}, \qquad VRF > 0 \tag{4.1}$$

where MSE is the smoothest estimator of variance. Secondly we define Importance Sampling factor (ISF) as being the ratio of the number of successes encountered during the simulation time. A success refers to a positive value of payoff function at the end of the replication time. Zero payoff is considered as a fail which just increases the MSE. We have one million replications for each input set. Every replication results in a success or fail. Simply ISF is calculated as

$$ISF = \frac{\sum Success_{ISMC}}{\sum Success_{cMC}}, \qquad ISF > 0 \tag{4.2}$$

We define ISF in order to explain the changes in VRF. Note that nominator and denominator subscripts are different for ISF and VRF. We keep this notation to make both key factors in the same direction. Thereby we can observe that increase in the VRF is provided by increase in the ISF. If we somehow get high variance reductions we can conclude that log-normal distribution doesn't fit well in specified regions. We tried to implement a number of underlying distributions as IS densities like Gama, truncated Pareto and mixture of log-normal distributions. We give the results for call and put options in Table 4.1 and Table 4.2, respectively. Numerical results indicate that it is possible to have high variance reductions for a wide range of moneyness.

Table 4.1: Black & Scholes premium, Monte Carlo simulation and IS Monte Carlo simulation results and VRF and ISF values are set in accordance for call options. The input set is {S={30,40,50,60,70}, K=50, r=10%, σ=20%, T=1 year}.

S	B&S	Monte Carlo Simulation			IS Monte Carlo				VRF	ISF
	Premium	Premium	Variance	StdDev	Premium	Variance	StdDev	IS density drift		
30	0.05384	0.05341	0.35151	0.59288	0.05377	0.00285	0.05334	r=0.62	123.55	58.53
40	1.39496	1.39636	12.87750	3.58852	1.39372	0.84678	0.92021	r=0.42	15.21	4.20
50	6.63484	6.63941	64.83510	8.05203	6.65658	7.16481	2.67672	r=0.29	9.05	1.53
60	15.12924	15.13415	133.34558	11.54754	15.12335	12.97372	3.60191	r=0.22	10.28	1.10
70	24.81572	24.82018	196.81735	14.02916	24.82817	12.41773	3.52388	r=0.19	15.85	1.02

Table 4.2: Black & Scholes premium, Monte Carlo simulation and IS Monte Carlo simulation results and VRF and ISF values are set in accordance for put options. The input set is {S={30,40,50,60,70}, K=50, r=10%, σ=20%, T=1 year}.

S	B&S	Monte Carlo Simulation			IS Monte Carlo				VRF	ISF
	Premium	Premium	Variance	StdDev	Premium	Variance	StdDev	IS density drift		
30	15.29571	15.30282	34.85894	5.90415	15.29456	5.66185	2.37946	r=0.01	6.16	1.02
40	6.63684	6.64039	33.89217	5.82170	6.63718	6.37729	2.52533	r=-0.07	5.31	1.31
50	1.87671	1.87079	12.26189	3.50170	1.87448	1.39664	1.18180	r=-0.20	8.78	2.90
60	0.37111	0.37388	2.26548	1.50515	0.37207	0.09837	0.31364	r=-0.33	23.03	10.44
70	0.05760	0.05862	0.32345	0.56873	0.05775	0.00347	0.05888	r=-0.45	93.30	49.65

We fixed four input parameters $T = 1$ year, $r = 10\%$, $\sigma = 10\%$, $K = 50$ and relaxed the spot price in the range of $S_0 \in [30{,}70]$ with unit increments. We developed an efficient C program that the simulation with one million replications takes just a few seconds. Figure 4.1a and Figure 4.1b show the graphs of key factors with respect to the spot price. Numerical results show that it is possible to have higher variance reductions for out-the money options. This implies that underlying stock dynamics can be represented better with an alternative IS distribution. Figure 4.2a and Figure 4.2b show key factors with respect to moneyness graphs for calls and puts, respectively.

Figure 4.1a Key factors are graphed with respect to spot price. All results are included for spot price in the range of $S_0 \in [30,70]$. More than 120 times of variance reduction is achieved for calls and 90 times for puts.

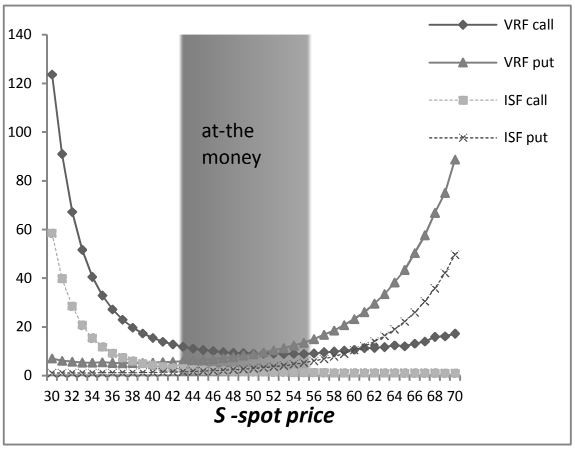
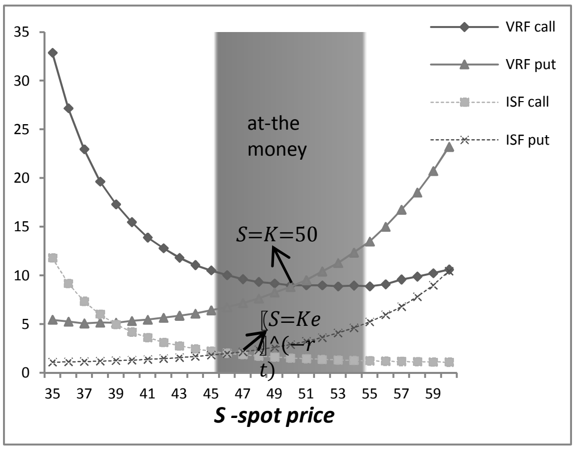


Figure 4.1b VRF curves for call and put pairs have an intersection point on the strike price value when $S = K = 50$. ISF curves has an intersection point at $S = Ke^{-rt}$. The two intersection points take place in at-the money region.



When simulation is repeated with different inputs, the results have similar trends. For example when strike price is $K = 40$ or $K = 60$ we have higher variance reductions for out-the money options and the intersection points always take place within at-the money region. Intuitively this suggests to have different underlying distributions for different moneyness regions. Then it would be possible to reflect the underlying price dynamics better. Finally we need to check the significance of the relation between VRF and ISF. We set regression test results in Table 4.3 which demonstrate strong relation for out-the money options and at-the money options.

Figure 4.2a Key factors-moneyness graph for call options

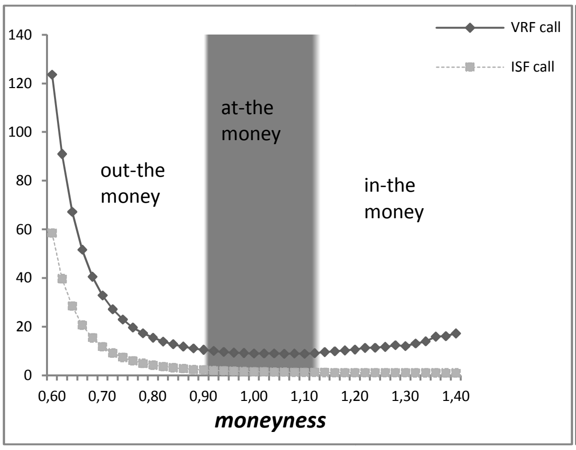
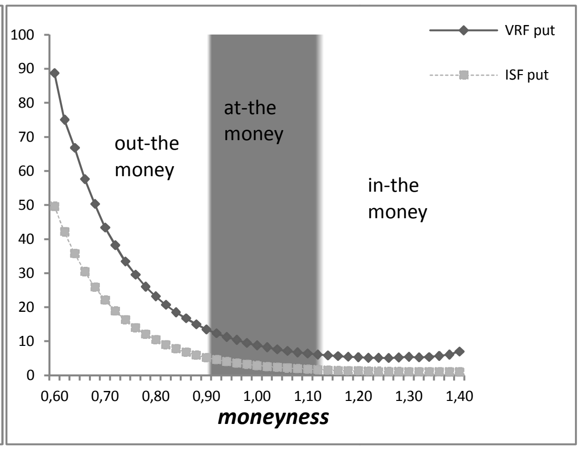


Figure 4.2b Key factors-moneyness graph for put options



Coefficient of determination, R^2 values are generally greater than 0.92 which shows a strong relation between VRF and ISF. On the other hand regression results do not reflect any relation for especially in-the money puts.

Table 4.3: Regression results for the key factors VRF and ISF. The comparison of R^2 values demonstrates strong relation for most of the option types except in-the money puts.

CALL			PUT	
SUMMARY OUTPUT		out-the money	SUMMARY OUTPUT	
Regression Statistics			*Regression Statistics*	
Multiple R	0.998867991		Multiple R	0.999533228
R Square	**0.997737264**		R Square	**0.999066673**
Adjusted R Square	0.997563207		Adjusted R Square	0.998994879
Standard Error	1.639265215		Standard Error	0.728023266
Observations	15		Observations	15
SUMMARY OUTPUT		at-the money	SUMMARY OUTPUT	
Regression Statistics			*Regression Statistics*	
Multiple R	0.960646777		Multiple R	0.999284893
R Square	**0.92284223**		R Square	**0.998570298**
Adjusted R Square	0.914269145		Adjusted R Square	0.998411443
Standard Error	0.158021041		Standard Error	0.093874356
Observations	11		Observations	11
SUMMARY OUTPUT		in-the money	SUMMARY OUTPUT	
Regression Statistics			*Regression Statistics*	
Multiple R	0.864050562		Multiple R	0.031463407
R Square	**0.746583374**		R Square	**0.000989946**
Adjusted R Square	0.727089787		Adjusted R Square	-0.075856981
Standard Error	1.310207471		Standard Error	0.52218281
Observations	15		Observations	15

5. CONCLUSIONS

We show insufficient aspects of the log-normal process in options pricing. We used simulation and variance reduction technique. The possibility of high variance reductions shows that original risk neutral measure of log-normal distribution cannot completely reflect the underlying price dynamics. Better alternatives could be found by easy combination of continuous distributions or second choice is regime switching models with a number of Lévy Processes whose parameters are dynamically changing. Both approaches could be better in the form of a risk neutral density for different moneyness regions. We used importance sampling in our analysis. The basic idea is to compute a correction factor to the importance sampling estimates.

With proper weights the correction factor compensates for statistical fluctuations. Hence output variance is decreased without disturbing the mean estimator. Numerical results indicate that it is better to have different underlying distributions for different moneyness regions. The contribution of our study is that we do not use empirical data which is always term dependent.

ACKNOWLEDGMENTS

The corresponding author would like to gratefully thank to TÜBİTAK for the support of BIDEB 2011.

REFERENCES

Aït-Sahalia, Y., Kimmel, R. (2006). Maximum likelihood estimation of stochastic volatility models. Journal of Financial Economics, 83 (2007) 413–452.

Bastani, A., Ahmadi, Z., Damircheli, D. (2013). A radial basis collocation method for pricing American options under regime-switching jump-diffusion models. Applied Numerical Mathematics, 65 (2013) 79–90.

Black, F., Scholes, M. (1973). The pricing of options and corporate liabilities. Journal of Political Economy, 81,637–659.

Boyle, P., Draviam, T. (2007). Pricing exotic options under regime switching. Insurance:, 40 (2007), 267-282.

Broadie, M., Glasserman, P. (1997). Pricing American style securities using simulation. Journal of Economic Dynamics and Control, 21, 1323-1352.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bonds and currency options. Review of Financial Studies, 6, 327-343.

Hull, J., White, A. (1987). The Pricing of Options on Assets with Stochastic. Journal of Finance, vol no 2, p 281-300.

Jackwert, C., Rubinstein, M. (1996). Recovering Probability Distributions from Option. Journal of Finance, 5 (1996), 1611-1631.

Liu, R., Zhao, J. (2013). A Lattice Method for Option Pricing with Two Underlying Assets in Regime-Switching Model. Journal of Computational and Applied Mathematics, S0377-0427(13)00081-2.

Longstaff, F. A. (1995). Option Pricing and Martingale Restriction. The review of Financial Studies, Vol. 8, sf. 1091-1124.

Merton, R. C. (1976). Options Pricing When Underlying Stock Returns are Discontinuous. Journal of Financial Economics, 3, 125-144.

Rubinstein, M. (1994). Implied binomial trees. Journal of Finance, 49, 771–818.

Stein, J. (1989). Overreactions in the Options Market. The Journal of Finance, 4, 1011-1023.

Yön, S. (2007). Importance Sampling for Queuing Simulations. Bogazici University Library, Master Thesis, Call # IE 2007 Y66.

Yön, S., Goldsman, D. (2006). Variance reduction via importance sampling. Istanbul Ticaret Universitesi Fen Bilimleri Dergisi, year 5 vol 10, p 35-41.