CZECH TECHNICAL UNIVERSITY IN PRAGUE
Faculty of Nuclear Sciences and Physical Engineering

# BACHELOR'S THESIS

2010 Jan Korbel

CZECH TECHNICAL UNIVERSITY IN PRAGUE
Faculty of Nuclear Sciences and Physical Engineering

# BACHELOR'S THESIS

## Application of Generalized Statistics in Econophysics


Author: Jan Korbel
Supervisor: Ing. Petr Jizba, PhD.

Year: 2010

## Acknowledgement

I would like to thank to my supervisor Ing. Petr Jizba, PhD., for all consultations and advices, he gave me during the last year. I have discovered a fascinating branch and I am very happy that he enabled me to get acquaint with that branch - econophysics. I would also like to thank to Jana Strouhalová for careful reading of my thesis and for corrections.

## Prohlášení

Prohlašuji, že jsem svou bakalářskou práci vypracoval samostatně a použil jsem pouze literaturu uvedenou v přiloženém seznamu.

Nemám závažný důvod proti užití školního díla ve smyslu §60 Zákona č 212/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

## Declaration

I declare, I wrote my bachelor thesis independently and exclusively with the use of cited bibliography.

I agree with the usage of this thesis in the purport of the §60 Act 121/2000 (Copyright Act).

V Praze dne ........................... ......................................

*Název práce:* **Aplikace zobecněné statistiky v ekonofyzice**
*Autor:* Jan Korbel
*Obor:* Matematické inženýrství
*Druh práce:* Bakalářská práce
*Vedoucí práce:* Ing. Petr Jizba, PhD.
Katedra fyziky, Fakulta jaderná a fyzikálně inženýrská

**Abstrakt**

Ekonofyzika je vědní obor, který analyzuje finanční trhy pomocí fyzikálních metod. Tato práce popisuje výhody použití Lévyho distribucí ve finančních trzích. Velké, otevřené systémy (jako finanční trhy) se dají často popsat distribucemi, které mají nekonečný první nebo druhý moment. Pro tato rozdělení neplatí centrální limitní teorém. Po zavedení nezbytných matematických pojmů je odvozena Black-Scholesova rovnice, která je nejběžněji používanou rovnicí pro oceňování opcí. Poslední část nastiňuje některá možná zobecnění Black-Scholesova modelu.

*Klíčová slova:* ekonofyzika, Black-Scholesova rovnice,
Wienerův proces, Itoův integrál, Lévyho rozdělení

*Title:* **Application of generalized statistics in econophysics**
*Author:* Jan Korbel

**Abstract**

Econophysics is a discipline that analyzes financial markets with physical methods. The thesis describes benefits of Lévy statistics applied to financial markets. Large, open systems (as financial markets) can be often well described by distributions that have infinite first or second moment. For these distributions central limit theorem does not hold. After introduction to necessary mathematical concepts, as stable distributions and stochastic calculus, we derive Black-Scholes equation, which is the most common equation used for option pricing. The last part outlines some possible generalizations of Black-Scholes model.

*Keywords:* econophysics, Black-Scholes equation,
Wiener process, Itō integral, Lévy distribution

# Contents

# Introduction

Every work about econophysics begins with citation of L. Bachelier's *Theory of speculation* [1]. We cannot start other way. His doctoral thesis was the first in that a physician tried to use mathematical and physical methods to financial markets. It seems to be the best definition of econophysics: *Physical methods used in financial markets.* Since that many physicists published papers about theory of financial risks and financial markets. We should mention at least Rosario Mantegna, Eugene Stanley [14], Steve Heston [12], Silvio Lévy, and Benoit Mandelbrot [13]. On the other hand, this branch is still quite new and there can be discovered many new ideas which would lead to better understanding of behavior of financial markets. There is much literature about econophysics, for all we can mention [2, 5, 16].

This thesis tries to show main benefits of econophysics. In financial markets are very important option-pricing models. For every stock and for traders, as well, is very important to have a mechanism, how to price a derivative of an asset. One possible and the most widespread model is Black-Scholes model. It was introduced by Fisher Black and Myron Scholes in 1973 and they received the Nobel prize for this model in 1997.

Before we derive Black-Scholes equation, we have to go through some mathematical background. The first part is about probability theory and it concentrates on the part of probability theory where is not possible to use central limit theorem, because distributions, we are interested in have infinite first or second moment. The second part is about stochastic processes, their definition and special classes of stochastic processes. Very important is Wiener process, a continuous analog of random walk. In the end of this chapter we define Itō stochastic integral that will be fundamental for deriving Black-Scholes equation. In the next part we make assumptions that are necessary to derive Black-Scholes equation and we derive it. The last part is devoted to limitations, generalizations and possible improvements of Black-Scholes model, especially about limitations of Wiener process. There will be shown some physical concepts used often in physics that can be suitable for financial markets, as well.

# Chapter 1

# Probability Theory

In the first chapter we begin with some basic formulas of probability theory often used in the next chapters, then go through random walk and show how works the central limit theorem and its limitations. The chapter will end with the generalization of central limit theorem and the concept of stale distributions. More about probability theory can be found in [9, 10].

## 1.1 Basic Results of Probability Theory

In this section we want to recapitulate some basic results of probability theory. These results are generally known, so we will not get to many details.

**Definition 1.1.1** *Let $\Omega$ be a sample space, $\mathcal{F}$ system of Borel sets ($\sigma$-algebra) on $\Omega$ and $\mu$ probability measure on $\mathcal{F}$, $(\Omega, \mathcal{F}, \mu)$ is a probability space. Every $\mathcal{F}$-measurable function on the probability space is called random variable.*

We usually assume random variables $X : \Omega \to \mathbb{R}(\mathbb{R}^n)$ and we will usually assume measure equivalent[1] to Lebesgue measure: $d\mu(x) = p(x)dx$. $p(x)$ is called probability density function and we will immediately see, how is defined.

**Definition 1.1.2** *For a random variable $X : \Omega \to \mathbb{R}$ is the function:*

$$F(x) = P[X(\omega) \leq x] \tag{1.1}$$

*called cumulative probability distribution (or cumulative density function - CDF).*

It is obvious that $F$ is positive nondecreasing function, i.e.: $F(-\infty) = 0, F(\infty) = 1$. $F(x)$ is a probability of $(-\infty, x]$. When $F(x)$ is differentiable function, we get definition of probability density function:

**Definition 1.1.3** *Let F(x) be a differentiable cumulative probability distribution. Then probability density function (PDF) p(x) is defined as:*

$$\mathrm{d}F(x) = p(x)\mathrm{d}x. \tag{1.2}$$

[1] equivalent means: for measures $\mu_1, \mu_2$: $\mu_1(X) = 0 \Leftrightarrow \mu_2(X) = 0$

Probability density function gives us probability in the interval $[x, x + dx]$. It is evident that probability in the interval $[a, b]$ is given by integral:

$$P[x \in [a, b]] = \int_a^b p(x)\mathrm{d}x. \tag{1.3}$$

We can define also the expectation value and variance as:

$$E(x) = \int_{-\infty}^{\infty} xp(x)\mathrm{d}x, \tag{1.4}$$

$$Var(x) = E\left((x - E(x))^2\right) = \int_{-\infty}^{\infty} (x - E(x))^2 p(x)\mathrm{d}x \tag{1.5}$$

when in discrete case integration changes to summation.

### 1.1.1 Multivariable Probability, Conditional Probability

Similarly we can define cumulative density for $n$ variables and $n$-variable probability density. For random variables $X_1, X_2, \dots, X_n$ is joint CDF given as:

$$F(x_1, x_2, \dots, x_n) = P[X_1 \leq x_1; X_2 \leq x_2; \dots; X_n \leq x_n], \tag{1.6}$$

and PDF is defined as:

$$\mathrm{d}F(x_1, x_2, \dots, x_n) = p(x_1, x_2, \dots, x_n)\mathrm{d}x_1\mathrm{d}x_2 \dots \mathrm{d}x_n. \tag{1.7}$$

**Definition 1.1.4** *Random variables $X_1, X_2$ are statistically independent, iff* [2]

$$p(x_1, x_2) = p'(x_1)p''(x_2). \tag{1.8}$$

When two variables are not statistically independent, we can define conditional probability. We take 2-dimensional joint distribution $p_2(x_1, x_2)$. We define marginal distribution $p_1(x_1)$:

$$p_1(x_1) = \int_{\mathbb{R}} p(x_1, x_2)\mathrm{d}x_1 \tag{1.9}$$

and similarly for $p_1(x_2)$.

**Definition 1.1.5** *Let $p_1(x_2) \neq 0$. Conditional probability is defined as:*

$$p(x_1|x_2) = \frac{p(x_1, x_2)}{p_1(x_2)}. \tag{1.10}$$

Conditional probability can be defined for more than one variables and more than one condition. The general formula has the form:

$$p_{n|m}(x_{m+1}, x_{m+2}, \dots, x_{m+n}|x_1, x_2, \dots, x_m) = \frac{p_{n+m}(x_1, x_2, \dots, x_{m+n})}{p_m(x_1, x_2, \dots, x_m)}. \tag{1.11}$$

**Note**: We shall denote n-point distribution $p_n$ and $n$-point distribution with $m$ conditions $p_{n|m}$.

[2] if and only if

For fixed $x_2$ is $p(x_1|x_2)$ density function for $x_1$, so e.g. conditional expectation value is:

$$E(x_1|x_2) = \int_{-\infty}^{\infty} x_1 p(x_1|x_2) \mathrm{d}x_1.$$

**Note**: If $x_1 \subset G$ then $E(x_1|G) = x_1$ a.s.[3]

We can define also covariance of two random variables. If $x, y$ are random variables, then covariance is defined as:

$$Cov(x,y) = E((x - E(x))(y - E(y))) = \int_{\mathbb{R}^2} (x - E(x))(y - E(y))p(x,y)\mathrm{d}x\mathrm{d}y, \tag{1.12}$$

we can define also correlation of two random variables:

$$Corr(x,y) = \frac{Cov(x,y)}{\sqrt{Var(x) \cdot Var(y)}}. \tag{1.13}$$

Covariance has this property:

$$Cov(x,y) = E((x - E(x))(y - E(y))) = E(xy) - E(x)E(y), \tag{1.14}$$

so for independent random variables is covariance and correlation zero. It is also obvious that:

$$Cov(x,x) = Var(x). \tag{1.15}$$

## 1.2 Random Walk

Let us start with the random walk on the line. The walker starts at time $t_0 = 0$ at position $x_0 = 0$. He can in the time $\Delta t$ make a step of the length $\Delta x$ with probability $p$ to the right and with the probability $q = 1 - p$ to the left. Our question is: What is the probability that the walker will be in time $t$ at position $x$?

Because we have a fixed time increment, we can work with discrete time steps labeled $n \in \mathbb{N}_0$. The same situation is with the position, we have a discrete net of points, they can be marked with $m \in \mathbb{Z}$. If the walker does a step to the right, his original position $m$ changes to $m + 1$ and analogically for step to the left. Probability $p(m,n)$ means probability in time $n$ at position $m$.

To calculate the probability we have to evaluate, how many ways we can get from the start to the position $m$ after $n$ steps. The walker makes $n$ steps together. If he ends at position $m$, he must do $m + l$ steps to the right and l steps to the left. From this we can calculate $l$ because $n = (m + l) + l$, so $l = \frac{n-m}{2}$. The walker makes $\frac{n+m}{2}$ steps to the left and $\frac{n-m}{2}$ steps to the right. The walker can do these steps in any order, so there are $\frac{n!}{(\frac{n+m}{2})!(\frac{n-m}{2})!}$ ways how to get to the point $m$.

Finally, the probability of doing $\frac{n+m}{2}$ steps to the left and $\frac{n-m}{2}$ to the right is: $p^{\frac{n+m}{2}}(1-p)^{\frac{n-m}{2}}$, so we can write that the probability is:

$$p(m,n) = \frac{n!}{(\frac{n+m}{2})!(\frac{n-m}{2})!} p^{\frac{n+m}{2}}(1-p)^{\frac{n-m}{2}}. \tag{1.16}$$

[3]a.s. : almost surely:$X \overset{a.s.}{=} Y \Leftrightarrow P[X = Y] = 1$

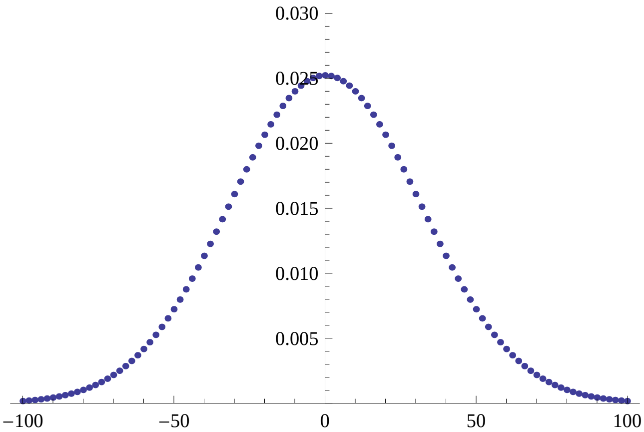


Figure 1.1: Random Walk distribution for $n = 1000$

From the last expression we can see that:

- For $|m| > n$ is $p(m, n) = 0$.
- If $n$ is even number, $m$ can be only even number and vice versa (otherwise the formula (1.16) would not have sense).
- The probability has *binomial* distribution with parameter $r = \frac{n+m}{2}$.

If we rewrite the probability distribution with terms of $r$, from the last property we can easily calculate the moments of the distribution by using power series:

$$E(r) = \sum_{r=0}^{n} rp(r,n) = \sum_{r=0}^{n} r\binom{n}{r}p^r(1-p)^{n-r} = \left[\sum_{r=0}^{n} r\binom{n}{r}p^r(1-p)^{n-r}x^r\right]_{x=1} =$$

$$\left[x\frac{d}{dx}\sum_{r=0}^{n}\binom{n}{r}p^r(1-p)^{n-r}x^r\right]_{x=1} = \left[nxp(px+q)^{n-1}\right]_{x=1} = np. \tag{1.17}$$

Similarly we calculate the second moment, while is:

$$E(r^2) = np + n(n-1)p^2,$$

then variance is:

$$Var(r) = E(r - E(r))^2 = E(r^2) - E(r)^2 = npq.$$

The deviation is:

$$\sigma = \sqrt{Var(r)} = \sqrt{npq}. \tag{1.18}$$

### 1.2.1 Relation to the Gaussian Distribution

Now, let us see, how can we approximate $p(m, n)$ for long times around the expected value $E(m)$.

We can write: $m = E(m) + \Delta m$.
Within we rewrite the terms included in the formula to these terms:

$$\frac{n+m}{2} = np + \frac{\Delta m}{2}, \quad \frac{n-m}{2} = nq - \frac{\Delta m}{2}.$$

We also use Stirling's formula

$$N! \approx \sqrt{2\pi N}\left(\frac{N}{e}\right)^N \tag{1.19}$$

to approximate factorial and calculate logarithm of probability distribution. We get:

$$\ln p(m,n) = (n+\frac{1}{2})\ln n - n - \frac{n+m+1}{2}\ln\frac{n+m}{2} + \frac{n+m}{2} +$$
$$\frac{n-m+1}{2}\ln\frac{n-m}{2} + \frac{n+m}{2} + \frac{n+m}{2}\ln p + \frac{n-m}{2}\ln q - \frac{1}{2}\ln 2\pi$$

$$= n\ln n + \frac{1}{2}\ln n - \left(Np + \frac{\Delta m}{2} + \frac{1}{2}\right)\ln\left(np + \frac{\Delta m}{2}\right) - \left(nq - \frac{\Delta m}{2} + \frac{1}{2}\right)\ln\left(nq - \frac{\Delta m}{2}\right)$$
$$+\left(np + \frac{\Delta m}{2}\right)\ln p + \left(nq - \frac{\Delta m}{2}\right)\ln q - \frac{1}{2}\ln 2\pi$$
$$= \frac{1}{2}\ln 2\pi npq - \left(np + \frac{\Delta m}{2} + \frac{1}{2}\right)\ln\left(1 + \frac{\Delta m}{2np}\right) - \left(nq - \frac{\Delta m}{2} + \frac{1}{2}\right)\ln\left(1 - \frac{\Delta m}{2nq}\right).$$

If we use an approximation of logarithm

$$\ln(1 \pm x) \approx \pm x - \frac{1}{2}x^2,$$

we get this formula:

$$\ln p(m,n) = \frac{1}{2}\ln(2\pi npq) - \frac{1}{2}\frac{(\Delta m)^2}{4npq} - \frac{\Delta m(q-p)}{4npq}.$$

Because $m$ is only a linear transform of $r$: $m = 2r - n$, variance of $m$ is proportional to variance of r. This leads to the fact that $\Delta m$ is proportional to the deviation of $r$. We get from (1.18) that the last term is proportional to the $n^{-\frac{1}{2}}$, so we can neglect it for large values of n. Finally, for large n we get this formula:

$$p(m,n) \to \frac{1}{\sqrt{2\pi npq}}\exp\left(-\frac{(\Delta m)^2}{8npq}\right), \quad \text{for } n \to \infty. \tag{1.20}$$

Because minimal $\Delta m$ for given $n$ is 2 (For odd $n$ the walker can be only in "odd" position and vice versa), we have to count with the factor 2 in the numerator.

We can recognize that the distribution converges to the normal distribution. If we calculate a sum of distributions $x_i$ where $x_i$ is distribution of one step of a walker [4], the sum of distribution goes to the normal distribution for $n \to \infty$.

[4]The distribution $x_i$ is an alternative distribution with probability $p$ for step to the right and $q = 1 - p$. to the left

This property describes central limit theorem, which we will discus in the next section.

We can now move to the continuous time and position. We can define time and position this way: $t = n\Delta t; \quad x = m\Delta x$.
Now we can send $\Delta t$ and $\Delta x$ to zero. When we require the condition:

$$2pq\frac{(\Delta x)^2}{\Delta t} = D = const,$$

then distribution becomes this form:

$$p(x,t)\mathrm{d}x = \frac{1}{\sqrt{4\pi Dt}}\exp\left(-\frac{(x-E(x))^2}{4Dt}\right)\mathrm{d}x. \tag{1.21}$$

## 1.3 Central Limit Theorem

In the last section we revealed an interesting property of random walk distribution. As the walker was doing more steps, the distribution was quite similar to the normal distribution. This is the property of many distributions, and the answer, which distributions converge to the normal distribution, gives us the central limit theorem:

**Theorem 1.3.1 (Central limit theorem)** *Let $X_i$ be an i.i.d.*[5] *variable with finite mean $\mu$ and finite variance $\sigma^2$. We define $Y_n = \sum_{i=1}^{n} X_i$.Then*

$$\frac{Y_n - \mu n}{\sigma\sqrt{n}} \xrightarrow{d} N(0,1) \tag{1.22}$$

*where $N(0,1)$ is normal distribution with $\mu = 0$ and $\sigma^2 = 1$ .*

**Notes:**

- A proof of CLT can be found in [10].
- Symbol: $\xrightarrow{d}$ means convergence in distribution. It is defined this way: *The sequence $X_n$ of random variables converges in distribution to X, iff*

  $$\lim_{n\to\infty} F(x_n) = F(x)$$

  where $F_n$ is CDF of $X_n$ and $F$ is CDF of $X$.
- CLT can be formulated as well for sequence of distributions that are not i.i.d. The only assumption is that every distribution $X_i$ must have finite mean and variance.

As we see, the fact that the first two moments must be finite is essential. Many distributions, that we know, satisfy these requirements, so they converge to the normal distribution. (We have seen that this property has binomial distribution, of course normal distribution and for example uniform distribution, as shown in figure 1.2.) On the other hand, in financial markets are important some distributions that are quite simply defined but have infinite first or second moment. For these distributions we cannot use central limit theorem. Next section will show us conditions when a sequence of distributions converges and what is the limit (or we can say attractive) distribution.

[5]independent, identically distributed

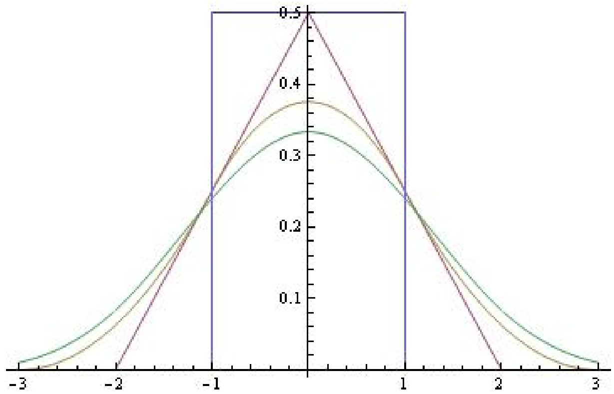


Figure 1.2: On the picture is illustrated the principle of the central limit theorem. We can see here the sum of $n$ independent random variables with uniform distribution ($p(x) = \frac{1}{b-a}$ for $a \leq x \leq b$, 0 else) Here are PDFs for $n = 1, 2, 3, 4$. Already for $n = 4$ it is possible to see that the distribution has a trend to get close to the normal distribution.

## 1.4 Stable Distributions

As we will see later, stable distributions have a huge role among other distributions. We will start first with definition:

**Definition 1.4.1** *Probability distribution $p(x)$ is stable, if it is invariant under convolution, i.e.: for every $a_1, a_2 > 0, b_1, b_2 \in \mathbb{R}$ exist $a > 0, b$, so that:*

$$p(a_1 l + b_1) * p(a_2 l + b_2) = \int_{-\infty}^{\infty} p\left(a_1[z-l] + b_1\right) p(a_1 l + b_1) \mathrm{d}l = p(az + b). \tag{1.23}$$

This definition is natural, because when we sum up two independent random variables with probability densities $p_1$ and $p_2$ respectively, we get the random variable with PDF that can be calculated by convolution. So we can say that stable distributions are distributions that do not change their form when we sum two variables with this probability distribution.

Quite useful is when we apply Fourier transform[6] to the equation (1.23). For $p(z) = f(l) * g(l)$ we can write:

$$p(k) = \mathcal{F}[p(z)] = \int_{-\infty}^{\infty} e^{-ikz} p(z) \mathrm{d}z = \int_{-\infty}^{\infty} e^{-ikz} \mathrm{d}z \int_{-\infty}^{\infty} f(z-l) g(l) \mathrm{d}l =$$

$$\left[\begin{array}{c} y = z - l \\ \mathrm{d}y = \mathrm{d}z \end{array}\right] = \int_{-\infty}^{\infty} e^{-ik(y+l)} \mathrm{d}y \int_{-\infty}^{\infty} f(y) g(l) \mathrm{d}l = f(k) g(k)$$

We have now a good criterion to decide whether a distribution is stable or not.

Let us give one example: Is normal distribution stable? For simplicity, we take normal distribution with $\mu = 0, \sigma^2 = 1$.

[6]Here we use the definition where in front of integral in Fourier transform is factor 1 and in inverse Fourier transform is factor $\frac{1}{2\pi}$, This definition does not preserve norm, but is a bit simpler in calculations.

$$\mathcal{F}\left[\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)\right] = \exp\left(-\frac{k^2}{2}\right)$$

So if we convolute two random variables $X_1, X_2$ with normal distribution, with Fourier transform we get:

$$\mathcal{F}\left[p(X_1) * p(X_2)\right] = \exp\left(-\frac{k^2}{2}\right)\,\exp\left(-\frac{k^2}{2}\right) = \exp\left(-k^2\right) = \mathcal{F}\left[\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{4}\right)\right].$$

That means that convolution of two normal distributions gives normal distribution, here with parameters $\mu = 0, \sigma^2 = 2$. The conclusion is that normal distribution is stable.

From this example is obvious that if we take for instance this Fourier image:

$$\mathcal{F}\left[p\right](k) = \exp\left(\frac{|k|}{2}\right),$$

then the convolution of these two distribution gives also the same distribution:

$$\exp\left(\frac{|k|}{2}\right)\exp\left(\frac{|k|}{2}\right) = \exp\left(|k|\right),$$

which is the Fourier transform of the Cauchy distribution:

$$\mathcal{F}^{-1}\left[\exp(-|k|)\right] = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-|k|}e^{ikx}\mathrm{d}k = \frac{1}{\pi}\frac{1}{1+x^2}. \tag{1.24}$$

We have two examples of stable distributions. Interesting about Cauchy distribution is the fact that it has indeterminate mean as well as variance. Why are stable distributions so important will explain the next theorem. It will answer us the question, what distributions can be the limit distribution of the sum of random variables. Let be $a_n \in \mathbb{R}, b_n > 0$ normalizing constants, $X_i$ random, i.i.d. variables with PDF $p(x)$. We define:

$$S_n = \frac{1}{b_n}\sum_{i=1}^{n} X_i - a_n \tag{1.25}$$

and we ask, what is the limiting PDF $L(x)$.

**Theorem 1.4.1 (Lévy, Kchintchin)** *A probability density $L(x)$ can be limiting distribution of sum (1.25) of independent, randomly distributed variables, only if $L(x)$ is stable.*

What more, next theorem gives us "explicit" form, of the distribution.

**Theorem 1.4.2** *A probability density $L_{\alpha\beta}(x)$ is stable, iff logarithm of its characteristic function has this form:*

$$\ln L_{\alpha\beta}(k) = ick - \gamma|k|^{\alpha}\left(1 + i\beta sgn(k)\omega(k,\alpha)\right) \tag{1.26}$$

*where:* $0 < \alpha \leq 2, \quad -1 \leq \beta \leq 1, \quad \gamma \geq 0, \quad c \in \mathbb{R},$

$$\omega(k,\alpha) = \begin{cases} \tan(\pi\alpha/2) & \text{if} \quad \alpha \neq 1 \\ \frac{2}{\pi}\ln|k| & \text{if} \quad \alpha = 1. \end{cases} \tag{1.27}$$

Both theorems and their proofs can be found in [11].

- Characteristic function of function $L_{\alpha\beta}(x)$ is its Fourier transform :

$$L_{\alpha\beta}(k) = \int_{-\infty}^{\infty} L_{\alpha\beta}(x) \exp(-ikx) \mathrm{d}x. \tag{1.28}$$

- Constant $c$ is the shift from origin and constant $\gamma$ is a scale parameter. These constants do not influence what kind of distribution it is. But constants $\alpha$ and $\beta$ influence shape and behavior of distribution.
- For some parameters $\alpha, \beta$ can happen that probability distribution $L_{\alpha\beta}(x)$ cannot be written in analytical form. It is the most common case. We will see a few exceptions.

Parameter $\alpha$ influences behavior for large $x$ and decides if the first or second moment exists. Parameter $\beta$ refers to asymmetry of the distribution. Note that the third moment (skewness) does not have to be defined. Distributions that satisfy condition (1.26) are called Lévy distributions.

For some special choices of parameters we get these properties:

- $\beta = 0$: $L_{\alpha\beta}$ is even function.
- $\alpha = 2$: $\omega(k, 2) = 0$ and it is normal distribution.
- $\alpha = 1, \beta = 0$: we get Cauchy distribution.
- $\beta = \pm 1$: in this case has a distribution an extreme asymmetry: for $\alpha < 1$ lies support of $L_{\alpha\beta}(x)$ in $(-\infty, c]$ for $\beta = 1$ and in $[c, \infty)$ for $\beta = -1$.

For us will be very important case when $\beta = 0$. Then the distribution is called *$\alpha$-stable Lévy distribution*. We shall look at Lévy distributions in the next section.

### 1.4.1 Lévy Distributions

As shown in the last section, Lévy distributions are distributions that satisfy condition (1.26). We assume $\alpha \leq 2$. (For $\alpha = 2$ is distribution Gaussian). In the next, we shall discus only distributions with $\beta = 0$. Very important attribute of Lévy distribution is behavior for $x \to \pm\infty$:

$$L_\alpha(x) \sim \frac{1}{|x|^{1+\alpha}} \quad \text{for} \quad |x| \to \infty. \tag{1.29}$$

We can show it without loss of generality for $c = 0$. Then the PDF would be calculated with inverse Fourier transform:

$$\begin{aligned} L_\alpha(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\gamma|k|^\alpha} e^{ikx} \mathrm{d}k = \frac{1}{2\pi} \int_0^{\infty} e^{-\gamma|k|^\alpha} \left(e^{ikx} + e^{-ikx}\right) \mathrm{d}k \\ &= \frac{1}{2\pi} \int_0^{\infty} e^{-\gamma|k|^\alpha} 2\Re\left(e^{ikx}\right) \mathrm{d}k = \frac{1}{\pi} \Re \int_0^{\infty} e^{-\gamma|k|^\alpha} e^{ikx} \mathrm{d}k. \end{aligned}$$

We will calculate the integral by the expansion of $e^{-\gamma|k|^\alpha}$ to the Taylor series:

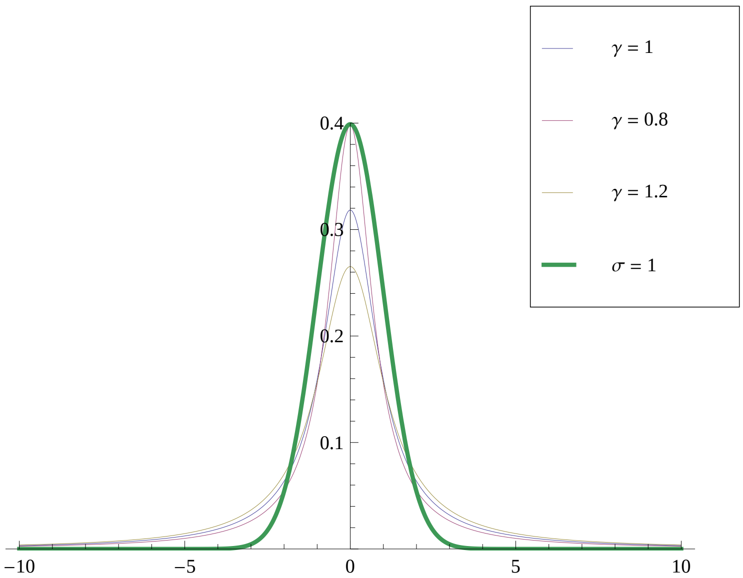


Figure 1.3: On the graph is shown Cauchy distribution with $\gamma = 1, \gamma = 0.8, \gamma = 1.2$. For comparison the thick graph is normal distribution with $\sigma = 1$.

$$\int_0^\infty e^{-\gamma|k|^\alpha} e^{ikx} \mathrm{d}k = \int_0^\infty \sum_{n=0}^{\infty} \frac{(-\gamma|k|^\alpha)^n}{n!} e^{ikx} \mathrm{d}k.$$

We change the order of summation and integration and after substitution we get the expression for the $\Gamma$ function:

$$\sum_{n=0}^{\infty} \frac{(-\gamma)^n}{n!} \int_0^\infty |k|^{\alpha n} e^{ikx} \mathrm{d}k = \sum_{n=0}^{\infty} \frac{(-\gamma)^n}{n!} \frac{\Gamma(\alpha n + 1)}{(-ix)^{\alpha n+1}}.$$

If we calculate the real part for the whole sum, and if we use identity:

$$\Re\left((\pm i)^{\alpha n+1}\right) = -\sin\frac{\pi\alpha n}{2},$$

we finally get asymptotic expansion of $L_\alpha(x)$:

$$L_\alpha(x) = -\frac{1}{\pi} \sum_{n=1}^{\infty} \frac{(-\gamma)^n}{n!} \frac{\Gamma(\alpha n + 1)}{|x|^{\alpha n+1}} \sin\frac{\pi\alpha n}{2}. \tag{1.30}$$

For large values (when $|x| \to \infty$) we can assume only the first term of the sum, so we get expression (1.29).

This attribute is called "heavy tails" or "fat tails", because for large $x$ tends the distribution to zero polynomially, but normal distribution tends to zero exponentially, for large $x$. From this law we can derive that Lévy distributions do not have variance, and for $\alpha \leq 1$ even mean does not exist. It is obvious that for finiteness of integral must integrand converge to zero for $x \to \pm\infty$.

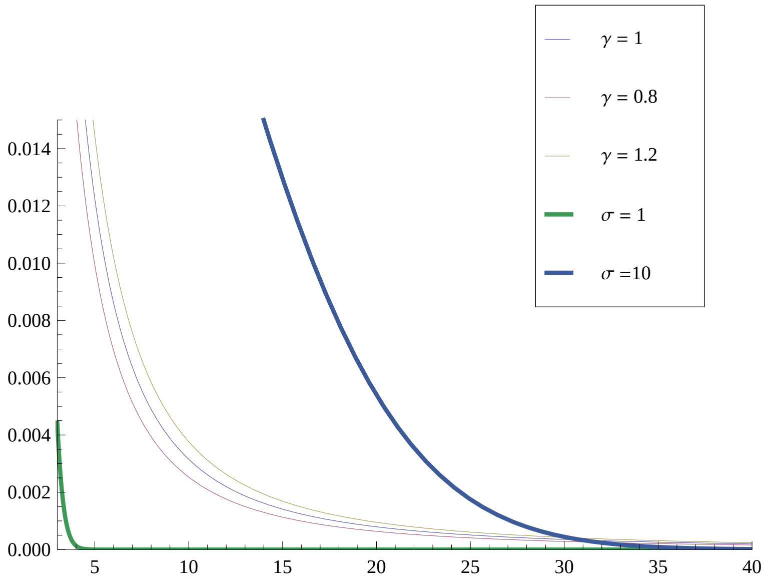


Figure 1.4: The graph shows "fat tails" of the Cauchy distribution with $\gamma = 1, \gamma = 0.8, \gamma = 1.2$. Here are the tails compared with normal distribution with $\sigma = 1, \sigma = 10$.

As said above, there is no algebraic expression of $L_\alpha(x)$ except for $\alpha = 1$, $\alpha = 2$. In comparison with Gaussian distribution has Lévy distribution sharper peak, and fatter tails. These distributions describe well situation when it is common to measure values around the expected value. But sometimes can come some large fluctuation and it is possible to measure values far from mid point. A good example are the financial markets, which will be discussed later.

# Chapter 2

# Stochastic Calculus

Stochastic calculus is a branch of mathematics working with random variables and stochastic processes. We will go through some special classes of stochastic processes and define the most important concept of the stochastic calculus - stochastic integral. More about stochastic calculus can be found e.g. in [6].

## 2.1 Stochastic Processes

In this section we define the term stochastic process, which is important for describing processes in financial markets. Basic term will be for us random variable, as defined in 1.1.1.

Typical example of random variable is coin tossing. Probability that head is thrown is $p$ and probability that tail is thrown is $q = 1 - p$ ($p \neq q$ for biased coin). It can be a situation from section 1.2 where can the walker throws a coin to decide, what will be his next step. So then we can say, if head is thrown, the walker goes to the right and if tail thrown, the walker goes to the left. The random variable can mean how the walker changes his position. The random variable $X$ can be

$$\begin{aligned} X(\text{head}) &= 1 \\ X(\text{tail}) &= -1. \end{aligned}$$

This random variable describes well one step of random walk. In many processes we can see that the process is composed of many random variables. Like random walk, where every step is random variable. These processes are called stochastic processes.

**Definition 2.1.1** *A stochastic process is a family of random variables $(\xi(t))_{t \in T}$, where $T \subset \mathbb{R}$ .*

If $T$ is discrete, we call it discrete stochastic process. An example is random walk. But often is $T = \mathbb{R}$ or $T = [0, \infty)$, typically, when parameter $t$ represents time.

**Definition 2.1.2** *Let $\xi(t, \omega)$, where $\omega \in \Omega$, be a stochastic process. Then function:*

$$t \mapsto \xi(t, \omega) \tag{2.1}$$

*is called sample path.*

An example is a particle doing a Brownian motion, where sample path is a possible path of the particle, or a stock index changing at time. If some phenomena is influenced by many negligible effects, like a particle of pollen grain in water when molecules of water hitting to the grain, or a stock index influenced by many trades, we can assume that the process is stochastic. For us will be important some special classes of stochastic processes.

### 2.1.1 Martingale Processes

Martingale processes are very important for financial applications. These processes have origin in gambling when gamblers wanted to know, if their odds to win are good or bad. Let us give this example. A gambler bets 1$ on coin tossing. The coin can be biased, so probabilities are not necessary $\frac{1}{2}$. The question is how much money will he win or lose on average after $n$ rounds. Before the first round the player has $s$$ and the probabilities are:

$$\begin{aligned} P[\text{win}] &= p, \\ P[\text{loose}] &= q = 1 - p. \end{aligned}$$

We have in chapter 1.2 calculated, what is average value of a walker, and here is it the same problem. From formula (1.17) we get that the walker does on average $np$ jumps to the right. Similarly, the average number of jumps to the left is $nq$. The average position of the walker and average winnings of the gambler) after $n$ steps is:

$$s + n(p - q). \tag{2.2}$$

It is clear that if $p > q$, then the game is for player favorable (he will get more and more many when playing more rounds) when $p < q$ is the game for a player disfavorable. When $p = q = \frac{1}{2}$, then the average winnings are 0, and the game is fair for player and for casino as well.

We can look, how is influenced the amount of money that the player has in $n$-th round with respect to the amounts of previous rounds. If the game is fair, the average amount must be the same as in previous round. This leads us to the definition of a martingale process.

**Definition 2.1.3** *The sequence of random variables $\{X_i\}_{i=1}^{\infty}$ is martingale iff:*

1. $E(|X_n|) < \infty$,
2. $E(X_{n+1}|X_1, X_2, \ldots, X_n) = X_n$.

I.e., martingale is a process, where the actual value of a stochastic process is the best for estimate next value.

If we take a sequence of random variables $\{X_i\}_{i=1}^{\infty}$ for which:

1. $E(X_1) = 0$,
2. $E(X_{n+1}|X_1, X_2, \ldots, X_n) = 0$,

then the sequence $\{Y_i\}_{i=1}^{\infty}$ defined as:

$$Y_n = K + X_1 + X_2 + \ldots + X_n \quad \forall K \in \mathbb{R}$$

is a martingale. Let us verify it:

$$E(Y_{n+1}|Y_1, Y_2, \ldots, Y_n) = E(Y_{n+1}|X_1, X_2, \ldots, X_n)$$

$$= E(K + X_1 + X_2 + \ldots + X_n + X_{n+1}|X_1, X_2, \ldots, X_n) =$$

$$= E(Y_n + X_{n+1}|X_1, X_2, \ldots, X_n) =$$

$$= E(Y_n|X_1, X_2, \ldots, X_n) + E(X_{n+1}|X_1, X_2, \ldots, X_n) = Y_n.$$

Process $\{X_i\}_{i=1}^{\infty}$ is called *fair game*, because it is the situation when average winnings/loss are 0.

An example of the martingale process can be given from our last example: if the probability to win (and to loose) is $\frac{1}{2}$, then the average winnings are 0. This process is martingale and we shall label it $X_n$. Another example of martingale is when we have process:

$$Y_n = X_n^2 - n. \tag{2.3}$$

Similar situation is with continuous stochastic process:

**Definition 2.1.4** *Continuous stochastic process* $\{\xi(t), t \in [0, \infty)\}$ *is called martingale iff:*

1. $\forall\ t : t \in [0, \infty)$ *is* $\xi(t)$ *integrable,*
2. $E(\xi(s)|\xi(r) : r \in [0, t]) = \xi(t)$ for $0 \leq s \leq t$.

Martingale processes are very important in markets, as we will see later, Wiener process defined in section 2.1.3 has martingale property. In next section will be described another property of stochastic processes.

### 2.1.2 Markov Processes

Markov processes are stochastic processes that have no memory. That means that the only information relevant for next evolution is the actual configuration of the system.
Let us have a stochastic process $\xi(t)$. If we denote:

$$p_{n|m}(x_1, t_1; t_2, x_2; \ldots; x_n, t_n|y_1, t_1; \ldots; y_m, t_m)$$

multivariable conditional probability, where $x_i, y_i$ belong to the random variable $\xi(t_i)$, then next definition gives us the way, how to write Markov property.

**Definition 2.1.5** *Stochastic process* $\xi(t)$ *has Markov property, iff for all* $t_1 < t_2 < \ldots < t_{n+1}$*:*

$$p_{1|n}(x_{n+1}, t_{n+1}|x_n, t_n; \ldots; x_1, t_1) = p_{1|1}(x_{n+1}, t_{n+1}|x_n, t_n). \tag{2.4}$$

We can recursively use this formula n point distribution $p_n$ a together with formula (1.11) we get:

$$p_n(x_1, t_1; \ldots; x_n, t_n) = p_1(x_1, t_1) \prod_{i=2}^{n} p_{1|1}(x_i, t_i|x_{i-1}, t_{i-1}). \tag{2.5}$$

That means that probabilities $p_1$ and $p_{1|1}$ give us full information about the process. When we use formula (2.5) for n=3:

$$p_3(x_1,t_1;x_2,t_2;x_3,t_3) = p_1(x_1,t_1)p_{1|1}(x_3,t_3|x_2,t_2)p_{1|1}(x_2,t_2|x_1,t_1)$$

and integrate over $x_2$, we get:

$$p_3(x_1,t_1;x_3,t_3) = p_1(x_1,t_1)\int_{\mathbb{R}} p_{1|1}(x_3,t_3|x_2,t_2)p_{1|1}(x_2,t_2|x_1,t_1)\mathrm{d}x_2.$$

When we divide this equation by $p(x_1,t_1)$, from (1.11) we get *Chapman-Kolmogorov* equation:

$$p_{1|1}(x_3,t_3|x_1,t_1) = \int_{\mathbb{R}} p_{1|1}(x_3,t_3|x_2,t_2)p_{1|1}(x_2,t_2|x_1,t_1)\mathrm{d}x_2. \tag{2.6}$$

Markov processes have essential role in Black-Scholes model. The assumption of no memory will lead us to using Itō calculus. But before that, we have to get acquainted with Wiener processes.

### 2.1.3 Wiener Process

In section 1.2 we were talking about random walk. At the end of the section, we tried to estimate, what would happen if time jumps tend to zero and the time begins to be continuous. This stochastic process is called Brownian motion and is well known from physics. A pollen grain is flowing in some liquid (e.g. water). The particle is much bigger, than liquid molecules, so we can see only the particle, not the molecules. As the molecules are moving, they hit to the particle. The energy of every hit is very small, but together it can move with the particle. Because we are not able to calculate motion equation for every molecule, we are not able to say, where will the particle move. But we want to know about the process something more, for example, what is the probability we find a particle in some area.
We would like to find a stochastic process that would have the same properties as brownian motion. For example, density function would have similar form as (1.21). Because, there is no preferred direction, the process should keep martingale property, as (symmetric) random walk.

**Definition 2.1.6** *Stochastic process W(t) is called Wiener iff:*

1. $W(0) \stackrel{a.s.}{=} 0$,
2. *sample paths* $t \mapsto W(t)$ *are continuous a.s.,*
3. $(\forall t,s|t>s) : W(t)-W(s) \sim N(0,t-s)$,
4. $W(t)$ *has independent increments of* $t$[1].

From this we can see that probability density function is:

$$p(x,t) = \frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{x^2}{2t}\right). \tag{2.7}$$

[1] $(\forall t_1,s_1;t_2,s_2)$: $W(t_1)-W(s_1)$ and $W(t_2)-W(s_2)$ are independent random variables

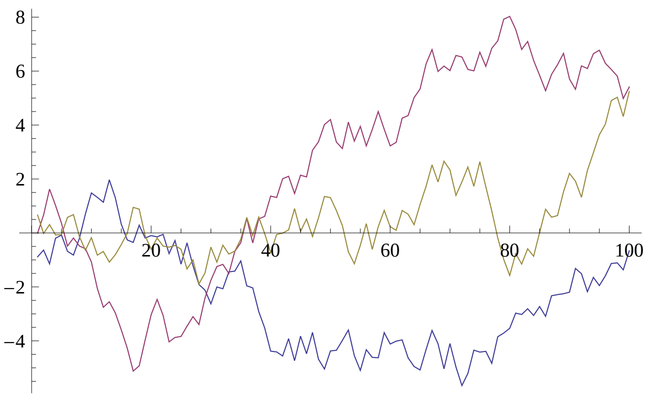

Figure 2.1: Typical paths of Wiener process

That is just equation (1.21) with $D = \frac{1}{2}$, so that is the same form, as we derived from random walk. The property 3 of definition of Wiener process tells us that Wiener process has Markov property.

We verify that $W$ has a.s. continuous sample paths: for $M > 0$

$$\lim_{\Delta t \to 0} P[|W(t+\Delta t) - W(t)| > M] = \lim_{\Delta t \to 0} \int_M^\infty \frac{2}{\sqrt{2\pi\Delta t}} \exp\left(-\frac{x^2}{2\Delta t}\right) \mathrm{d}x =$$

$$= \int_M^\infty \lim_{\Delta t \to 0} \frac{2}{\sqrt{2\pi\Delta t}} \exp\left(-\frac{x^2}{2\Delta t}\right) \mathrm{d}x = \int_M^\infty 2\delta(x)\mathrm{d}x \ = 0.$$

On the other hand, sample paths are a.s. nowhere differentiable:

$$\lim_{\Delta t \to 0} P\left[\left|\frac{W(t+\Delta t) - W(t)}{\Delta t}\right| > M\right] = \lim_{\Delta t \to 0} \int_{M\Delta t}^\infty \frac{2}{\sqrt{2\pi\Delta t}} \exp\left(-\frac{x^2}{2\Delta t}\right) \mathrm{d}x =$$

$$\int_{-\infty}^\infty \underbrace{\lim_{\Delta t \to 0} \frac{1}{\sqrt{2\pi\Delta t}} \exp\left(-\frac{x^2}{2\Delta t}\right)}_{\delta(x)} \mathrm{d}x - \underbrace{\lim_{\Delta t \to 0} \int_0^{M\Delta t} \frac{2}{\sqrt{2\pi\Delta t}} \exp\left(-\frac{x^2}{2\Delta t}\right) \mathrm{d}x}_{\to 0} = 1.$$

Very important property of Wiener process is that $W(t)$ is martingale:

$$E(W(s)|W(r), r \in [0,t)) = E(W(s) - W(t)|W(r), [0,t)) + E(W(t)|W(r), [0,t))$$

$$= E(W(s) - W(t)) + W(t) = W(t).$$

Also process $W(t)^2 - t$ is martingale:

$$E(W(s)^2|W(r), r \in [0,t)) = E((W(s) - W(t))^2|W(r)) + E(2W(t)W(s)|W(r)) -$$

$$-E((W(t))^2|W(r)) = E((W(s) - W(t))^2) + 2W(t)E(W(s)|W(r)) - W(t)^2 = s - t + W(t)^2.$$

Another interesting and in the next section needful result is expected value of joint distribution of Wiener process in time $t, s$. In case $s > t$:

$$E(W(s)W(t)) = \int\int xyp(s-t, x-y)p(t,y)\mathrm{d}y\mathrm{d}x =$$

$$\int\int xy\frac{1}{\sqrt{2\pi(s-t)}}\exp\left[-\frac{(x-y)^2}{2(s-t)}\right]\frac{1}{\sqrt{2\pi t}}\exp\left[-\frac{y^2}{2t}\right]\mathrm{d}x\mathrm{d}y = t. \tag{2.8}$$

So generally:

$$E(W(s)W(t)) = min\{s,t\}. \tag{2.9}$$

With knowledge of Wiener process, we can deduce Itō stochastic integral, which will be the next section about.

## 2.2 Itō Stochastic Integral

Let us imagine a stochastic process whose increments are proportional to Wiener process

$$dX(t) = \sigma dW(t). \tag{2.10}$$

This description has a few unfavorable attributes: at first the probability that $X(t) < 0$ is positive for any $t > 0$. Imagine that $X(t)$ would be a price of an asset. This price can be obviously never negative. It is usual to express total gain or total loss in proportion of all the money invested (as we will see in the next chapter). We can rewrite (2.10) as:

$$\frac{dX(t)}{X(t)} = \sigma dW(t). \tag{2.11}$$

But there is another problem: as shown in previous section, $W(t)$ is a.s. nowhere differentiable and $dW(t)$ is not a common differential. The way, how to interpret these equations gave Kiyoshi Itō and is known as Itō integral. We shall follow the definition of Riemann-Stieltjes integral.

**Definition 2.2.1** *Let us consider function f(t) over interval $(t_0, t)$, partition of this interval, i.e. a sequence of points $\{t_i\}_{i=0}^{n}$, where $t_0 < t_1 < \ldots < t_{n-1} < t_n = t$ and sequence of $\tau_i$, where $\tau_i \in [t_{i-1}, t_i]$. Then stochastic integral is equal to:*

$$\int_{t_0}^{t} f(t')\mathrm{d}W(t') = \mathrm{ms-}\lim_{n\to\infty}\sum_{i=1}^{n} f(\tau_i)\left[W(t_i) - W(t_{i-1})\right], \tag{2.12}$$

*where* ms−lim *is mean square convergence, i.e.:*

$$\mathrm{ms-}\lim_{n\to\infty} X_n = X \Leftrightarrow \lim_{n\to\infty} E((X_n - X)^2) = 0.$$

**Note:** As in case of Riemann integral, where we can define Riemann-Stieltjes integral, we can here define more general form of stochastic integral. Let us have for example continuous function F(x), then we can define stochastic integral:

$$\int_{t_0}^{t} g(t')\mathrm{d}[F(W(t'))] = \mathrm{ms-}\lim_{n\to\infty}\sum_{i=1}^{n} g(\tau_i)F\left[W(t_i) - W(t_{i-1})\right]. \tag{2.13}$$

Contrary to the Riemann integral, the value of stochastic integral depends on the choice of $\tau_i$, because $W(t)$ is a.s. nowhere differentiable. For every interval $[t_i, t_{i+1}]$ we choose one point and we enumerate, how integral depends on the choice of these points.

We take $f(t) = W(t)$, $\tau_i = \alpha t_{i-1} + (1-\alpha)t_i, \alpha \in (0,1)$ and calculate:

$$E\left(\sum_{i=1}^{n} W(\tau_i)\left[W(t_i) - W(t_{i-1})\right]\right) =$$

$$= \sum_{i=1}^{n} E[W(\alpha t_{i-1} + (1-\alpha)t_i)W(t_i)] - E[W(\alpha t_{i-1} + (1-\alpha)t_i)W(t_{i-1})] =$$

$$= \sum_{i=1}^{n} (t_i - \alpha t_{i-1} + (1-\alpha)t_i) = \alpha(t_n - t_0). \tag{2.14}$$

We used formula (2.9) to simplify the sum.

Here are two most common and most important choices of $\alpha$:

- $\alpha = 0$ defines Itō integral,
- $\alpha = \frac{1}{2}$ defines Stratonovich integral.

We will be using Itō integral. We will not be using any other kind of stochastic integral, so we will not distinguish it from general stochastic integral.

Stratonovich integral is widely used in other parts of physics, especially in quantum field theory. Its advantage is that things like integration per partes, or substitution in integral are similar to Riemann integral, contrary to Itō integral. On the other hand, using Itō integral has a few advantages. At first, Itō integral is natural description of processes, like evolution of stock price when from present value we estimate the future vaules. The second reason is that stochastic process defined by integral:

$$X(t) = \int_0^t f(t')\mathrm{d}W(t') \tag{2.15}$$

is martingale.
If we denote: $\Delta W_i = W(t_i) - W(t_{i-1})$ and without loss of generality we assume that for every division of interval $(0,t)$ $\exists k \ : \ t_k = s$, then:

$$E(X(t)|X(s)) = \text{ms}-\lim_{n\to\infty}\left(\sum_{i=1}^{n} E(f(t_{i-1})\Delta W_i|X(s))\right),$$

for $i < k$ $(t_i < s)$:

$$E(f(t_{i-1})\Delta W_i|X(s)) = f(t_{i-1})\Delta W_i,$$

for $i \geq k$ $(t_i \geq s)$ is $W(t_i)$ independent of $W(t_j)$ $\forall j \leq k$:

$$E(f(t_{i-1})\Delta W_i|X(s)) = f(t_{i-1})E(\Delta_i W) = 0.$$

The integral is the equal to:

$$\text{ms}-\lim_{k\to\infty}\left(\sum_{i=1}^{k} E(f(t_{i-1})\Delta W_i)\right) = X(s). \tag{2.16}$$

In stochastic calculus has special role class of functions, which do not anticipate the future behavior of Wiener process.

**Definition 2.2.2** *Function $g(t)$ is called non-anticipating of $t$ when $g(t)$ is statistically independent of $W(s) - W(t)$ for all s,t; $s > t$.*

Examples of non-anticipating functions are: $W(t)$ , $\int_{t_0}^{t} \mathrm{d}W(t)$ or $\int_{t_0}^{t} g(t)\mathrm{d}W(t)$ when $g(t)$ is non-anticipating. Let $g(t)$ be a non-anticipating function, then these stochastic integrals can be calculated this way:

$$\int_{t_0}^{t} g(t')\mathrm{d}\left[W(t')^2\right] = \int_{t_0}^{t} g(t')dt', \tag{2.17}$$

$$\int_{t_0}^{t} g(t')\mathrm{d}\left[W(t')^n\right] = 0 \text{ for } n \geq 3. \tag{2.18}$$

We shall prove the first formula, the second one can be proved analogically. With: $\Delta W_i = W(t_i) - W(t_{i-1})$, $\Delta t_i = t_i - t_{i-1}$, $g(t_i) = g_i$, we can prove formula (2.17) from definition:

$$\lim_{n\to\infty} E\left[\left(\sum_{i=1}^{n} g_{i-1}\left((\Delta W_i)^2 - \Delta t_i\right)\right)^2\right] =$$

$$= \lim_{n\to\infty} E\left[\sum_{i=1}^{n} g_{i-1}^2((\Delta W_i)^2 - \Delta t_i)^2\right.$$

$$\left.+2\sum_{j=1}^{n}\sum_{i=j}^{n} g_{i-1}g_{j-1}((\Delta W_j)^2 - \Delta t_j)((\Delta W_i)^2 - \Delta t_i)\right]$$

$$= \lim_{n\to\infty} \left(\sum_{i=1}^{n} E[g_{i-1}^2]E\left[((\Delta W_i)^2 - \Delta t_i)^2\right]\right.$$

$$\left.+2\sum_{i=1}^{n}\sum_{j=i}^{n} E\left[g_{i-1}g_{j-1}((\Delta W_j)^2 - \Delta t_j)\right] E\left[((\Delta W_i)^2 - \Delta t_i)^2\right]\right)$$

from the fact that $g(t)$ and $W(t)$ are non-anticipating.

Because $E\left[(\Delta W_i)^2\right] = \Delta t_i$ (from property 3 of definition 2.1.6),

$$E\left[((\Delta W_i)^2 - \Delta t_i)^2\right] = 2(\Delta t_i)^2. \tag{2.19}$$

We can rewrite the last expression as:

$$\lim_{n\to\infty} 2\sum_{i=1}^{n}\left(E(g_{i-1}^2) + 4\sum_{j=i}^{n} E(g_{i-1}g_{j-1})(\Delta t_j)^2\right)(\Delta t_i)^2 \to 0$$

when e.g. g is bounded. Similarly we would be able to derive formula (2.18).

### 2.2.1 Stochastic Differential Equations

Now, with knowledge of stochastic integrals, we can interpret equation (2.11).

**Definition 2.2.3** *A stochastic process x(t) obeys (Itō) stochastic differential equation (SDE):*

$$\mathrm{d}x(t) = a(x(t), t)\mathrm{d}t + b(x(t), t)\mathrm{d}W(t) \tag{2.20}$$

*iff for all* $t_0, t$*:*

$$x(t) = x(t_0) + \int_{t_0}^{t} a(x(t'), t')\mathrm{d}t' + \int_{t_0}^{t} b(x(t'), t')\mathrm{d}W(t'). \tag{2.21}$$

Now equation (2.18) is well defined stochastic equation in form:

$$dX(t) = \sigma X(t)dW(t). \tag{2.22}$$

With this terminology we can rewrite formulas (2.17) and (2.18) to simply form:

$$\mathrm{d}[W(t)^2] = dt, \tag{2.23}$$

$$\mathrm{d}[W(t)^n] = 0, \tag{2.24}$$

which is only the shorter notation of expressions written above.

### 2.2.2 Properties of Itō Integral, Itō's Lemma

Now, we would like to calculate some Itō integrals and stochastic differential equations. For this we would like to learn, how to deal with stochastic integrals. Some features we already know:

- linearity

$$\int_{t_0}^{t} (f + \alpha g)(t')\mathrm{d}W(t') = \int_{t_0}^{t} f(t')\mathrm{d}W(t') + \alpha \int_{t_0}^{t} g(t')\mathrm{d}W(t'), \tag{2.25}$$

- martingale property

$$X(t) = \int_{0}^{t} f(t')\mathrm{d}W(t') \ : \ E(X(t)|X(r) : r \in (0, s)) = X(s). \tag{2.26}$$

As in Riemann calculus, we would like to have some more tricks, how to deal with integrals, like integration per partes, or substitution in integral case, respectively Leibnitz rule or chain rule in differential case. For us will be important to have some form of chain rule, which is essential for deriving Black-Scholes equation in the next chapter.

We shall assume a non-anticipating function $F(t, x) \in C^1[\,[0, \infty)] \times C^2[\mathbb{R}]$, and consider division $\{t_i\}_{i=1}^n$ of interval $[0, t]$,where:

$$0 = t_0 < t_1 < \ldots < t_{n-1} < t_n = t,$$

and sequences $\{\tau_i\}_{i=1}^n$,$\{\xi_i\}_{i=1}^n$ , where $\tau_i, \xi_i \in (t_{i-1}, t_i)$. We shall denote:

$$\frac{\partial F(t, x)}{\partial t} = F_t(t, x) \ , \ \frac{\partial F(t, x)}{\partial x} = F_x(t, x).$$

Then:

$$F(t,W(t)) - F(0,W(0)) = \sum_{i=1}^{n} F(t_i,W(t_i)) - F(t_{i-1},W(t_{i-1})) =$$

$$= \sum_{i=1}^{n} F(t_i,W(t_i)) - F(t_{i-1},W(t_i)) + \sum_{i=1}^{n} F(t_{i-1},W(t_i)) - F(t_{i-1},W(t_{i-1})),$$

and we shall use Taylor formula to the first order in $(\Delta t_i)$. But in the second sum is not possible to expand the expression only to the first order, because we know from (2.23), (2.24) that even the terms with $(\Delta W_i)^2$ will contribute to the terms of $(\Delta t_i)$.

With Lagrange formula and Taylor polynomial we can rewrite last expression as:

$$\sum_{i=1}^{n} F_t(\tau_i,W(t_i))\Delta t_i + \sum_{i=1}^{n} F_x(t_{i-1},W(t_{i-1}))\Delta W_i + \frac{1}{2}\sum_{i=1}^{n} F_{xx}(t_{i-1},W(\xi_i))(\Delta W_i)^2 =$$

$$= \sum_{i=1}^{n} F_t(\tau_i,W(t_i))\Delta t_i + \frac{1}{2}\sum_{i=1}^{n} F_{xx}(t_{i-1},W(t_{i-1}))\Delta t_i +$$

$$+ \sum_{i=1}^{n} F_x(t_{i-1},W(t_{i-1}))\Delta W_i + \frac{1}{2}\sum_{i=1}^{n} F_{xx}(t_{i-1},W(t_{i-1}))((\Delta W_i)^2 - \Delta t_i) +$$

$$+ \frac{1}{2}\sum_{i=1}^{n} [F_{xx}(t_{i-1},W(t_{i-1})) - F_{xx}(t_{i-1},W(\xi_i))](\Delta W_i)^2.$$

Now, we use to the expression mean square limit and look what is going to happen to each term.

- Because $F_t$ and $F_{xx}$ are continuous on $[t_0,t]$, and $W(t)$ has a.s. continuous sample paths, then the first converges a.s. to the integral:

$$\int_{t_0}^{t} F_t(t,W(t))\mathrm{d}t \tag{2.27}$$

  and the second sum converges a.s. to the integral:

$$\frac{1}{2}\int_{t_0}^{t} F_{xx}(t,W(t))\mathrm{d}t. \tag{2.28}$$

- because $F_x$ is differentiable (and bounded) on $[t_0,t]$, and W(t) is again a.s. continuous, the third sum converges (in the sense of mean square convergence) to Itō integral:

$$\int_{t_0}^{t} F_x(t,W(t))\mathrm{d}W(t). \tag{2.29}$$

- The fourth sum converges to zero in mean square convergence, because with help of (2.19):

$$E\left[\left(\frac{1}{2}\sum_{i=1}^{n}F_{xx}(t_{i-1},W(t_{i-1}))((\Delta W_i)^2-\Delta t_i)\right)^2\right]=$$

$$=\frac{1}{2}\sum_{i=1}^{n}E[F_{xx}(t_{i-1},W(t_{i-1}))^2]E[((\Delta W_i)^2-\Delta t_i)^2]\to 0. \tag{2.30}$$

- Because the same arguments of the continuity, the last sum converges to zero:

$$\frac{1}{2}\sum_{i=1}^{n}[F_{xx}(t_{i-1},W(t_{i-1}))-F_{xx}(t_{i-1},W(\xi_i))](\Delta W_i)^2\leq$$

$$\leq\underbrace{\sup_i[F_{xx}(t_{i-1},W(t_{i-1}))-F_{xx}(t_{i-1},W(\xi_i))]}_{\to 0}\underbrace{\sum_{i=1}^{n}(\Delta W_i)^2}_{\to(t-t_0)}\to 0. \tag{2.31}$$

From (2.27) - (2.31) we get formula:

$$F(t,W(t))-F(0,W(0))=\int_{t_0}^{t}\left(F_t(t,W(t))+\frac{1}{2}F_{xx}(t,W(t))\right)\mathrm{d}t+$$
$$+\int_{t_0}^{t}F_x(t,W(t))\mathrm{d}W(t). \tag{2.32}$$

If we rewrite the statement to the differential formulation, we get special version of Itō's lemma:

$$\mathrm{d}F(t,W(t))=[F_t(t,W(t))+\frac{1}{2}F_{xx}(t,W(t))]\mathrm{d}t+F_x(t,W(t))\mathrm{d}W(t). \tag{2.33}$$

If we compare it with "classical" chain rule:

$$\mathrm{d}F(t,y(t))=F_t(t,y(t))\mathrm{d}t+F_y(t,y(t))\mathrm{d}y(t) \tag{2.34}$$

we see that there is one term added. That is caused by the fact that W(t) is stochastic process, so it is called Itō correction.

If we have a general stochastic process $x(t)$ fulfilling stochastic differential equation:

$$\mathrm{d}x(t)=a(t)\mathrm{d}t+b(t)\mathrm{d}W(t), \tag{2.35}$$

we can formulate general version of Itō's lemma:

**Theorem 2.2.1** *Let F(t,x) be real function from $C^1[[0,\infty)]\times C^2[\mathbb{R}]$ and x(t) stochastic process fulfilling equation (3.4), then:*

$$\mathrm{d}F(t,x(t))=\left(F_t(t,x(t))+a(t)F_x(t,x(t))+\frac{1}{2}b(t)^2F_{xx}(t,x(t))\right)\mathrm{d}t$$
$$+b(t)F_x(t,x(t))\mathrm{d}W(t). \tag{2.36}$$

With these instruments we are able to derive Black-Scholes equation, which is next chapter about.

# Chapter 3

# Black-Scholes Model

In this chapter we would like to present Black-Scholes model usually used for pricing of European options. Before we derive Black-Scholes equation, we have to learn some terminology of financial markets.

## 3.1 Financial Markets

Financial market is a place, where money are invested in commodities, lend or borrowed, invested, etc. Traders can deal in commodities as wheat, gold or oil, securities as assets, bonds or options, or foreign currencies. Whereas commodities are real goods, securities are just papers that represent some financial value. An example of security is an asset. Asset represents part of wealth of e.g. some corporation.

There are two ways how to trade: on the stock exchange, or by the over-the-counter exchange. The stock exchange is official place for trading these financial instruments. There are given rules how to trade, and the stock organizes the process of trading. On the other hand over-the-counter (OTC) exchange means that the traders arrange the conditions of trade among themselves. Trading at the stock exchange is more common. The best known stock exchange is New York Stock Exchange (NYSE) in the USA, Tokyo Stock Exchange (TSE) in Japan and London Stock Exchange (LSE) in the UK. The exchange gives to traders all possible information about traded asset and determines the price of assets due to actual supply and demand. Stock exchanges guarantee that certain rules are obeyed, e.g. only approved assets can be traded, the traders are tested if they can pay their obligations, etc.

### 3.1.1 Financial Derivatives: Futures and Options

In the following we will be interested in the special class of securities that are called *derivatives*. The name comes from the fact that their value is derived from the value of other, more basic asset. This asset is called *underlying*. Mathematically said: if $S(t)$ is price of an underlying in a time $t$, then the price of a derivative is only a function of $S(t)$. There are two most important kinds of derivatives: futures and options. In the case of *future*, buyer and seller sign a contract, which states that after exactly defined period the owner of an asset sells this asset for an arranged price to the buyer. In this context is often used a terminology that buyer is in long position or shortly long and the seller is in short position or simply short.

What is the reason to trade in futures? Good reason for trading in futures is to minimize the risk. Let us give an example: A farmer grows wheat. He has to pay for the fuel for tractors, for the pesticides etc. But in the market the price of wheat fluctuates because of supply and demand of wheat. When the wheat will be ready to crop, the price of wheat can be high and the farmer will earn money, but it can happen that the price will be low and he will loose money. In order to minimize the risk, he decides to make a contract that he will sell the corn for a given price. For him is it favorable, because he will be able to pay for the fuel and for the seeds for next season. For the party that is in long position it can be profitable when the price of corn will be higher than agreed price. The future carries the risk to the long position.

While in case of future both parties participating in the contract have the same duties (the short has to sell the asset and the long has to buy the asset), in case of *options* is the situation different. The simplest kind of options are European options. The seller of the option is called *writer* and guarantees to the buyer of the option, who is called *holder*, the right (not the duty) to buy or sell the asset in given time (also called expiration time) for given price. In case of the right of buying asset from the writer is the option called *call option*, in case of selling to the writer, the option is called *put option*. There exist other types of options, as American option that can be exercised any time before its expiration time. It would be more difficult to describe these option in mathematical formulation of our problem, so we will not work with them.

The price of the option is given by stock exchange in case of exchange options, or by the parties agreeing the contract, as in case of OTC options. We can ask, how to define a price of an option, so that it would be optimal to both writer and holder. Possible solution gives Black Scholes model.

## 3.2 Assumptions of the Black-Scholes Model

As said above, Black-Scholes model is option pricing model used specifically for European options. In order to derive some solutions we have to make a few assumptions, to simplify the model that we would be able to formulate with mathematical structures we already know.

In the next we denote $T$ as the expiration time of the option (= the time when option can be exercised), $S(t)$ price of an asset and $O(S,t)$ price of a derived option.

### Possible Risks on the Market

In finance we recognize a few kinds of risk. The first is *credit risk*. This kind of risk comes from the possibility that the other side of contract would not be able to pay its claims. This is inhibited in case of exchange options, where the stock exchange guarantees you to pay for the debtor. The other risk is *operational risk*. This is the kind of risk when a firm looses because of its internal policy, behavior of its employees or systems. For example in computer program is hidden a bug, and the program gives wrong predictions, so the firm looses money. Or when the firm forgets to pay a bill and must pay some penalty.

We will assume that there is no credit and operational risk. The only risk assumed is *market risk* that comes from trading assets and their unfavorable evolution.

### Efficient Market Hypothesis

We will assume that the market is *efficient*. That means that:

1. It is possible to get immediately all information relevant to trading.
2. The market is *liquid*.
3. There is no, or negligible *market friction*.

With term liquid we mean that prices of products reflect current situation and it is possible to buy or sell a product in any time without any delay. Market friction is term for all trading costs, as provisions, taxes etc. We assume that the sum of these costs is negligible to the sum of money invested in the market.

These assumptions have one important consequence: If the market is efficient, then all information about market is contained in the present state of market. This leads to the conclusion that processes on the market must be Markov processes.

### Geometric Brownian Motion

Next question is, how to describe an asset price evolution. From here we will denote the price of an asset in time $t$ as $S(t)$. Let us imagine the situation that we have some amount of money and we want to invest them. One possibility is to deposit the money in the bank. At the beginning we have sum of $M(0)$ in the bank. The bank guarantees a risk-free interest rate r for the deposit. If the bank would pay an interest once at the end of period $[0, T]$, then

$$M(T) = M(0)\,(1 + rT)\,.$$

If the bank would pay twice, at first at the half of the period and then at the end of time period, then:

$$M(T) = M(0)\left(1 + \frac{rT}{2}\right)^2,$$

etc. For $n \to \infty$ would it look like this:

$$M(T) = \lim_{n\to\infty} M(0)\left(1 + \frac{rT}{n}\right)^n = M(0)\exp(rT). \tag{3.1}$$

This interest is called continuously compounded. If we rewrite it in differential notation, then

$$\mathrm{d}M(t) = rM(t)\mathrm{d}t. \tag{3.2}$$

Then we have another choice how to invest our money. We can invest in some assets. The evolution of their price would have two parts, deterministic and stochastic part. The first will describe general trends of the asset price, the second will describe fluctuations on the market. The first part will be similar to the equation (3.2) so it will be:

$$\mathrm{d}S_{det}(t) = \mu S(t)\mathrm{d}t, \tag{3.3}$$

the parameter $\mu$ is called *drift parameter* and we can assume $\mu > r$, because otherwise would be not wise to buy asset when the risk-free interest rate is higher.

We assume that the stochastic part is stochastic process that should have Markov property. This fulfills for example Wiener process, so:

$$\mathrm{d}S_{stoch}(t) = \sigma S(t)\mathrm{d}W(t). \tag{3.4}$$

We will assume $\sigma > 0$, the factor $\sigma$ is called *volatility.* Processes with higher volatility are not so stable and their price can more fluctuate. Given (3.3) and (3.4) together we get a stochastic differential equation describing price of an asset:

$$\mathrm{d}S(t) = \mu S(t)\mathrm{d}t + \sigma S(t)\mathrm{d}W(t) \tag{3.5}$$

with initial condition that the price of asset in the time 0 is positive,i.e.: $S(0) > 0$.

We assume that interest rate $r$ and volatility $\sigma$ are constants. On the other hand there are models that assume that volatility or interest rate are not constant in time. We will mention it in the next chapter.

Now we can solve this stochastic differential equation. When we use Itō's lemma for the function $\ln[S(t)]$, we get:

$$\mathrm{d}\ln[S(t)] = \left(\mu - \frac{1}{2}\sigma^2\right)\mathrm{d}t + \sigma\mathrm{d}W(t). \tag{3.6}$$

Because

$$\int_0^t \mathrm{d}W(t') = W(t),$$

by interpreting SDE we get:

$$\ln S(t) = \ln S(0) + (\mu - \frac{1}{2}\sigma^2)t + \sigma W(t), \tag{3.7}$$

and written in integral form:

$$S(t) = S(0)\exp\left[(\mu - \frac{1}{2}\sigma^2)t + \sigma W(t)\right]. \tag{3.8}$$

This process is called geometric Brownian motion, or sometimes exponential martingale. This process will be for us the process that simulates stock price movement at time. We see that $\ln S$ is normally distributed, and we can calculate mean and variance. Because from property:

$$E(W(t)) = 0, \quad Var(W(t)) = t$$

we get:

$$E(\ln[S(t)/S(0)]) = (\mu - \frac{1}{2}\sigma^2)t, \quad Var(\ln[S(t)/S(0)]) = \sigma^2 t.$$

We can finally write a PDF of a stochastic process $S(t)$:

$$p(S,t) = \frac{1}{S\sqrt{2\pi\sigma^2 t}}\exp\left(-\frac{\left(ln[S/S(0)] - [\mu - \frac{1}{2}\sigma^2]t\right)^2}{2\sigma^2 t}\right). \tag{3.9}$$

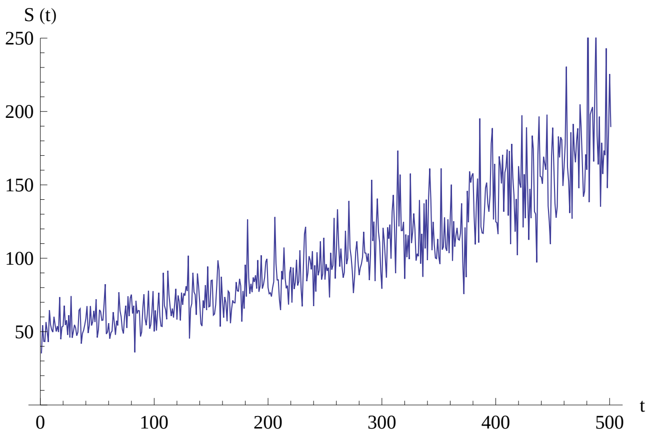

Figure 3.1: Geometric Brownian motion

**Other Assumptions**

There are some more assumptions important for the model:

1. The first is that the underlying pays no dividends.
2. The second is that the underlying is infinitely divisible, i.e. the holder can own $k$ assets, where $k$ is not necessarily the natural number (k can be any real number).
3. The last is that there exists no *arbitrage*.

The term arbitrage means, if a trader lends money from the bank with interest rate r, then does not exist any risk-less strategy (e.g. deposit these money to other bank, etc.) with drift parameter $\nu$, so that $\nu > r$. That would lead to the situation that if borrowed some amount of money $C$, after time T the investor would have:

$$C\exp(\nu T) - C\exp(rT) > 0$$

and he would get richer and richer even if he had no money at the begin.

## 3.3 Delta-Hedging Strategy

Now, we look at the position of the writer of option. He runs a risk of selling the underlying asset under the the market price. This risk should be compensated by the price of the option that he sold. We would like to find an optimal strategy that would eliminate the risk. There are many risk hedging strategies. We will not need to go through all of them, for more about hedging, see [5, 14].

One possible strategy is called *delta-hedging*. During the time before expiration time can writer own only a fraction of an underlying (it is possible from assumptions above) depending on price $S(t)$. When $S$ is rising, the writer should own a bigger fraction of an underlying, because it is more probable that the option will be exercised, and vice versa. The writer of course needs money for enlarging the fraction of asset when necessary. These money can be invested in the risk-free process, e.g. deposited to the bank. The process has to be risk-free, because the writer must be sure that he would be able to buy the part of asset in any case. So if we denote $\Delta(t)$ the fraction of the underlying owned, and $\Pi(t)$ the sum of money, necessary to possible buys, then the total richness $R(t)$ of the writer should be:

$$R(t) = \Delta(t)S(t) + \Pi(t). \tag{3.10}$$

How can the writer get this sum of money? He can get the money right from selling the option, so the writer should require that the price of the option is equal to the sum of money he needs. That means:

$$O(t) = R(t) = \Delta(t)S(t) + \Pi(t). \tag{3.11}$$

If we make a differential of $O(t)$, we get:

$$\mathrm{d}O(t) = \Delta(t)\mathrm{d}S(t) + r\Pi(t)\mathrm{d}t, \tag{3.12}$$

where we assumed $\mathrm{d}\Delta(t) = 0$. This we can perceive that asset price (and from that option price, because it is function of $S(t)$) is not influenced by change of $\Delta(t)$, On the contrary, $\Delta(t)$ is influenced by fluctuations of asset price. The second term comes from the fact that the best risk-free investment has interest rate $r$, and from (3.2).

## 3.4 Black-Scholes Equation

If we rewrite differential $\mathrm{d}O$ with help of Itō's lemma, we get:

$$\mathrm{d}O(S(t),t) = \left(\frac{\partial O}{\partial t} + \mu S(t)\frac{\partial O}{\partial S} + \frac{1}{2}\sigma^2 S(t)^2 \frac{\partial^2 O}{\partial S^2}\right)\mathrm{d}t + \sigma S(t)\frac{\partial O}{\partial S}\,\mathrm{d}W(t), \tag{3.13}$$

$$\mathrm{d}O(S(t),t) = \left(\frac{\partial O}{\partial t} + \frac{1}{2}\sigma^2 S(t)^2 \frac{\partial^2 O}{\partial S^2}\right)\mathrm{d}t + \frac{\partial O}{\partial S}\,\mathrm{d}S(t). \tag{3.14}$$

If we compare the coefficients in equations (3.12) and (3.14), we get these two formulas:

$$\Delta(t) = \frac{\partial O}{\partial S}, \tag{3.15}$$

$$r\Pi(t) = \frac{\partial O}{\partial t} + \frac{1}{2}\sigma^2 S(t)^2 \frac{\partial^2 O}{\partial S^2}. \tag{3.16}$$

The first equation comes from the delta-hedging, and tells us that if the writer will own in every time this amount of underlying, he will eliminate the risk coming from the option. Sometimes can be this equation interpreted as a definition of delta-hedging strategy.

If we rewrite the second equation with help of equation (3.10), we get Black-Scholes formula for European options:

$$\frac{\partial O}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 O}{\partial S^2} + rS\frac{\partial O}{\partial S} - rO = 0. \tag{3.17}$$

Note that in this equation occurs no $\mu$. The price of option is independent of drift parameter of an underlying.

### 3.4.1 Boundary Conditions

It would seem, that Black-Scholes model describes all options with the same way, no matter, if they are call options, or put options. The difference between these two kinds is hidden in boundary conditions of Black-Scholes equation. In the next, we shall denote $T$ as expiration time, $K$ as *strike price* (= agreed price, for that can be the option exercised), $C(S,t)$ will be price of a call option and $P(S,t)$ will be price of a put option.

For call and put options we have these boundary conditions:

**call option:**

$$C(S,T) = \max(S(T) - K, 0) \tag{3.18}$$

$$C(0,t) = 0 \tag{3.19}$$

$$C(S,t) \to S \quad \text{for } S \to \infty, \tag{3.20}$$

**put option:**

$$P(S,T) = \max(K - S(T), 0) \tag{3.21}$$

$$P(0,t) = K \exp(-r(T-t)) \tag{3.22}$$

$$P(S,t) \to 0 \quad \text{for } S \to \infty. \tag{3.23}$$

We will look at the price of the call option in the time $T$. If the price of an asset $S(T)$ would be lower than strike price, the holder would not exercise the option, because for him is cheaper to buy an asset directly. So $C(S,T) = 0$. If the price of an asset would be higher than strike price, the writer will require that he would like at least not to loose any money. So the option should cost at least the difference between actual market price and strike price. Because if the price would be more, than this minimum price, the holder will not buy the option. We get for both cases, $S(T) < K, S(T) > K$, the condition (3.18).

Very similar situation is with the put option. If the asset price $K$ would be higher than strike price at expiration time $T$, the holder would not exercise the option, so $P(S,T) = 0$. If $S(t) < K$, the holder would exercise the option, and the writer would loose $K - S(T)$. This loss would writer compensate with the price of the option, so similarly given these two conditions together, we get (3.21).

Another conditions we get, if we look at behavior of option price for $S(t) = 0$. Then we can see from (3.5) that if at any time $t_0$ would be price of an asset 0, then would stay 0 for all times.

If we look at call option when a price of an underlying asset is 0, than it is obvious that this option will not be exercised, so the price of call option is 0. On the other hand, the put option will be certainly exercised. So the price of put asset should be strike price, discounted by risk free interest rate.

The last condition results from behavior when the price of an asset goes to infinity. Then we can assume $S \gg K$, so the price will be influenced only by price $S$.

### 3.4.2 Solution of the Black-Scholes Equation

Using Green's function method, we get the solution of the Black-Scholes equation (the whole solution and all details are in [16]):

$$C(S,t) = S\Phi(d_1) - K\exp(-r(T-t))\Phi(d_2) \tag{3.24}$$

$$P(S,t) = K\exp(-r(T-t)) - S + C(S,t), \tag{3.25}$$

where

$$d_1 = \frac{\ln(\frac{S}{K}) + (r + \frac{\sigma^2}{2})(T-t)}{\sqrt{\sigma}(T-t)} \quad d_2 = d_1 - \sqrt{\sigma}(T-t)$$

and $\Phi(x)$ is CDF of normal distribution.

This solution has some advantages, but as we will see in the next chapter, the Black-Scholes model has some limitations. The next chapter will be all about improvements of Black-Scholes model.

# Chapter 4

# Beyond Black-Scholes Model

The Black-Scholes model is the most used model for pricing options. Its advantage is that the solution is in analytical form, and the fact that the model is simple and easy to understand. Moreover, it really works well for the most of time. On the other hand, the model has some assumptions that sometimes make the it unreal. We will look at the limitations of Black-Scholes model and outline some of possible generalizations.

## 4.1 Limitations of Black-Scholes Model

In order to derive Black-Scholes equation, we had to make a few assumptions. Many of them are in usual situations a good approximation of real markets. We will analyze them.

The first sort of deflections of the model have technical character. The assumption of no dividends is not necessary, if we modify the function for the asset price. If we know, how will be dividends paid (the sum of dividends paid in time t $D(t)$ is known function od time), then

$$S'(t) = S(t) \exp\left( - \int_t^T D(t')\mathrm{d}t' \right). \tag{4.1}$$

As well the assumption of constant $\sigma$ and $r$ can be relaxed if they are known functions of time. Then we can use average values:

$$\sigma^2 = \frac{1}{t-T} \int_t^T \sigma^2(t')\mathrm{d}t, \qquad r = \frac{1}{t-T} \int_t^T r(t')\mathrm{d}t. \tag{4.2}$$

Assumptions as divisibility of assets or no market friction assumption are usually well fulfilled in bigger transactions. Discussions about validity of efficient market hypothesis is beyond this thesis. These problems are more economical character and will not discus them.

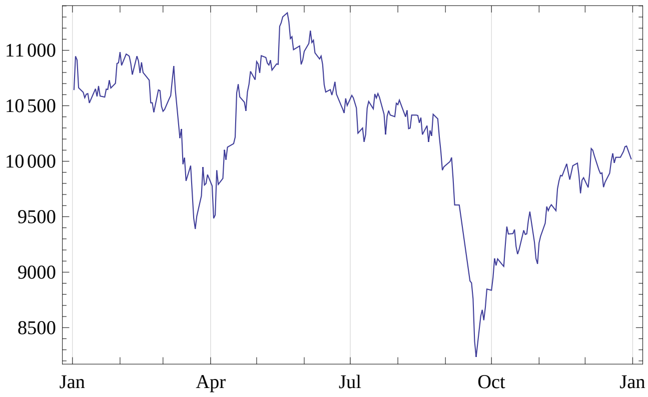


Figure 4.1: Figure shows the evolution of *Dow Jones Industrial Averages* in the year 2001. Dow Jones index consists of 30 companies as Coca-Cola, IBM, etc. On the graph we can see rapid fall after terrorists attacks in 11. 9. 2001. The index steeply declined after the attacks and after a few weeks it began to rise again.

### 4.1.1 Large Deviations

For us will be more important the problem of suitable model for an asset price. We have decided that the part of price evolution, which describes stochastic behavior of an asset price, will be described by the Wiener process. Wiener process is described by normal distribution (or the process of price evolution is described by log-normal distribution) and for this distribution are very improbable large deviations.

We can return again to the continuous random walk which is described with the the the standard deviation $\sigma$. We can ask, what is the probability that the walker will get further during time interval $\Delta t = 1$, than for example $\sigma, 2\sigma, 3\sigma \ldots$ For the distance $\beta > 0$ from actual position is the probability given by CDF of normal distribution $\Phi$ :

$$P[|\Delta x| > \beta] = 2(1 - \Phi(\beta)) \tag{4.3}$$

For value $\sigma$ is the probability 31.73%, for $2\sigma$ is the probability 4.55% and for example, the probability that the the walker will be further than $6\sigma$ is about $1.97 \cdot 10^{-7}$%. That means that large steps are very improbable. Because prices of assets are described with distributions based on normal distribution, for them it means that for the prices are large fluctuations almost impossible, or they can happen very rarely, something like once per $10^7$ units of time when talking about fluctuations $6\sigma$ large. But we know that in financial markets are these large fluctuations are more frequent. It is possible to see on figures 4.1, 4.2 and 4.3, that events as natural disasters, terrorists attacks, or fatal decisions of politicians can influence markets so much that even very improbable fall or rise can happen. So during the period when markets are not much influenced by these outer effects, can markets be well described by Wiener process. When markets are influence by e.g. a natural disaster, then Wiener process is not good for modeling asset prices. It is the same when we imagine a grain of pollen in the container with water. When the container is at the rest, the process of grain

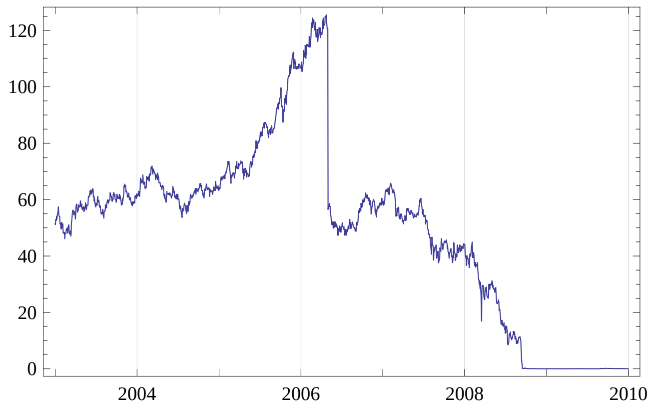


Figure 4.2: On the graph we can see the evolution of *Lehman Brothers Holding, Inc.* stock. The investment bank suffered from subprime mortgage crisis in 2006 - 2007 and this led the bank to the bankruptcy in 15. 8. 2008. This incident started the global financial crisis.

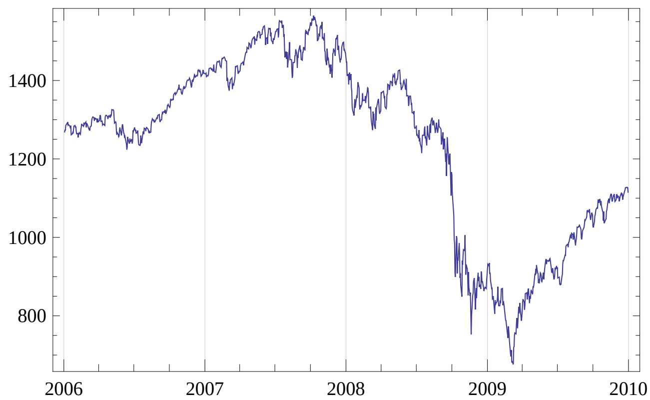


Figure 4.3: This graph shows the fall of *Standard & Poor's 500 index* during global financial crisis in the years 2008 and 2009. During the year the index fell from 1,565.15 points in 9. 10. 2007 to 676.53 points in 9. 3. 2009. The fastest slump was at the beginning of financial crisis in the September 2008 when the index suddenly lost more than 300 points.

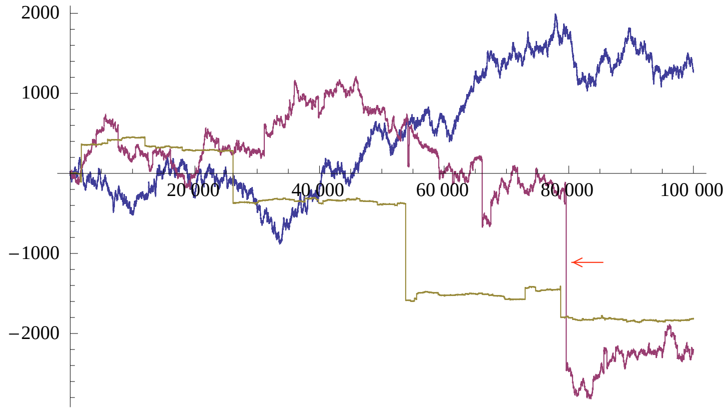


Figure 4.4: Figure shows a simulation of typical random walk with different probability distribution of steps. In the simulation were made $10^5$ steps. In the first case (blue line) is it classical Wiener process, i.e. the probability distribution is normal with mean 0 and variance 5. In the second case (purple line) was the probability distribution Student distribution with parameter $\nu = 1.7$. Student distribution with this parameter has mean $\mu = 0$ and indefinite variance, so the behavior of the "walker" is quite similar to situation when the distribution would be Lévy with parameter $1 < \alpha < 2$, and Lévy distribution is its limit distribution. The arrow marks the large deviation, this behavior we can see in financial markets, e.g. during some disaster. In the third case (yellow line) is the probability distribution is Cauchy distribution. This distribution is too sharp and has too fat tails, so it products very small fluctuations alternated by huge jumps.

movement can be described by Wiener process, but when we hit the container, then the process cannot be well described by Gaussian process.

We can ask, what distributions are the best for describing financial markets. It may be surprising that heavy tailed distributions can be a good candidate for the suitable distribution. It may seem that because they may not have defined mean or variance, they cannot describe any real physical system, less so financial markets. But it is not so. Heavy tails property leads to the fact that large deviations are much more probable, than in normal distribution case. On the other hand, they usually have sharper peak around the mean (if it is defined), so these processes stay for some time oscillate around the expected value and then make a big jump and then again.

Special role have Lévy distributions, because they are attractors of sums of random variables. For longer times, we can see that the process that is described by heavy tail distribution behaves similarly as the process with Lévy distribution with appropriate parameter $\alpha$. For example for S&P index is $\alpha \simeq 1.4$ [16, p. 162].

In the next section we would like to suggest some financial models that would work better even for extreme situations, and Lévy distribution will usually have a fundamental role in these models.

## 4.2 Generalizations of Black-Scholes Model

Here we would like to show a few models that extend Black-Scholes model. Our aim is not to precisely derive all details, but we would like to suggest, how is possible to improve ideas shown in previous chapter.

### 4.2.1 Double Stochastic Equation

The idea of double stochastic equation comes from Steven Heston and is published in [12]. The main concept is simple: assumption of constant volatility, or volatility that we can predict is usually not satisfied. We do not usually know, how volatile will markets be. The idea is that not only asset price, but also volatility is a stochastic process. If volatility fulfills Ornstein-Uhlenbeck process [15], we get these two equations:

$$\mathrm{d}S(t) = \mu S(t)\mathrm{d}t + \sigma(t)S(t)\mathrm{d}W_1(t) \tag{4.4}$$

$$\mathrm{d}\sigma(t) = -\beta\sigma(t)\mathrm{d}t + \delta\mathrm{d}W_2(t). \tag{4.5}$$

Processes $W_1(t)$ and $W_2(t)$ are not independent, but

$$Corr\left(W_1(t), W_2(t)\right) = \rho, \tag{4.6}$$

where $|\rho| < 1$.

Using Itō's lemma we get an expression for differential of $\sigma^2(t)$, which will be more useful for next calculations:

$$\mathrm{d}\sigma^2(t) = [2\delta^2 - \beta\sigma^2(t)]\mathrm{d}t + 2\delta\sigma(t)\mathrm{d}W_2(t). \tag{4.7}$$

This process is called square-root process and usually is written in this form:

$$\mathrm{d}v(t) = \kappa[\theta - v(t)]\mathrm{d}t + \xi\sqrt{v(t)}\mathrm{d}W_2(t). \tag{4.8}$$

We will be also using this form, so $\sigma^2(t) = v(t)$.

For option price $O$ we assume that $O = O(S, v, t)$ and calculate $\mathrm{d}O$ with Itō's lemma for more stochastic processes [7]:

$$\mathrm{d}F(t, x_1, \ldots, x_n) = \frac{\partial F}{\partial t}\,\mathrm{d}t + \sum_{i=1}^{n}\frac{\partial F}{\partial x_i}\,\mathrm{d}x_i + \frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial^2 F}{\partial x_i \partial x_j}\,\mathrm{d}[Cov(x_i, x_j)]. \tag{4.9}$$

For option price we get (for lucidity we omit arguments):

$$\begin{aligned}\mathrm{d}O = \Bigg(&\frac{\partial O}{\partial t} + \mu S\frac{\partial O}{\partial S} - \kappa[\theta - v]\frac{\partial O}{\partial v} + \frac{1}{2}vS^2\frac{\partial^2 O}{\partial S^2} + \frac{1}{2}\xi^2 v\frac{\partial^2 O}{\partial v^2} + \\ &+ \xi S v\rho\frac{\partial^2 O}{\partial S\partial v}\Bigg)\mathrm{d}t + \sqrt{v}S\frac{\partial O}{\partial S}\,\mathrm{d}W_1(t) + \xi\sqrt{v}\frac{\partial O}{\partial v}\,\mathrm{d}W_2(t).\end{aligned} \tag{4.10}$$

We again assume sum of money needful to suitable hedging strategy:

$$R(t) = \Delta(t)S(t) + \Pi(t) + \gamma(t)C(S, v, t), \tag{4.11}$$

where we should own not only the part of underlying, but also a part of some concrete option.

If we calculate a differential od $R(t)$ and compare coefficients similar to Black-Scholes model, we get *Double-stochastic equation* [12]:

$$\frac{\partial O}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 O}{\partial S^2} + \rho\xi vS\frac{\partial^2 O}{\partial S\partial v} + \frac{1}{2}\xi^2 v\frac{\partial^2 O}{\partial v^2} + \\ + rS\frac{\partial O}{\partial S} + (\kappa[\theta - v] - \lambda(S, v, t))\frac{\partial O}{\partial v} - rO = 0. \tag{4.12}$$

Term $\lambda$ represents market price of volatility risk and depends on the concrete market. The other advantage is that this model is applicable for more kinds of options because of suitable choice of $\lambda$.

Next sections will be about concepts that have their background in thermodynamics. We will see that assets dealt in financial markets and molecules of gas moving in a container can have some similar properties and it can be useful to use these methods in financial markets.

### 4.2.2 Superstatistics

The main idea of superstatistics is: "*Superstatistics* is statistics of statistics." One of the first scientist engaged in superstatistics was Christian Beck [4]. The first stimulus came from thermodynamics. We know that energy distribution of molecules of gas with inverse temperature $\beta$ in equilibrium is:

$$p(E, \beta) = \frac{1}{Z(\beta)}\rho(E)e^{-\beta E}, \tag{4.13}$$

where $Z(\beta)$ is partition function and $\rho(E)$ is density of states with energy $E$.

Here we assume that the system is in equilibrium. That means that temperature in the whole system is the same. We can assume that it is true locally, i.e. we can divide the volume into many small areas with constant temperature. We can count, how many of these areas have given temperature. If the volume of these areas were infinitely small, we can measure the distribution of the temperature $f(\beta)$. Then PDF of energy, $p(E)$ is given by superposition of these density functions:

$$p(E) = \int_0^\infty f(\beta)\frac{1}{Z(\beta)}\rho(E)e^{-\beta E}\mathrm{d}\beta. \tag{4.14}$$

In our case can volatility change during the time. A probability distribution of asset price is given by (3.9), we can assume that volatility is constant in every time interval $(t_i, t_{i+1})$, but it is not necessary, to be the same constant in all of time intervals. If these intervals were infinitely small, volatility could be given by suitable distribution. The question is, what is the right distribution of volatility.

We can see that the distribution of $X = \ln S$ is normal, so it depends on volatility as:

$$p(X, \sigma) \sim \frac{1}{\sqrt{\sigma^2}}\exp\left[-\frac{X^2}{\sigma^2}\right].$$

If the PDF depends on the parameter $\beta$ more like $\frac{1}{\beta}$, then it is natural to use for probability distribution of $\beta$ inverse gamma distribution[1] (more in [3, 4]). From the form of PDF, we can recognize that the role of "inverse temperature" here has the term $\frac{1}{\sigma^2}$. If we put these distributions into formula (4.14) now for $\sigma^2$ we get new distribution $\tilde{p}$ of $X$. These distributions are heavy-tailed, we will demonstrate in on an example that is analytically expressible.

As distribution $f(\sigma^2)$ we take inverse gamma distribution $InvGamma(\alpha, \beta)$ and we assume that $\mu = \frac{1}{2}\sigma^2$. Then the distribution is:

$$\tilde{p}(X,t) = \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)\sigma^{2\alpha+2}\sqrt{2\pi\sigma^2 t}} \exp\left(-\frac{\beta}{\sigma^2} - \frac{X^2}{2\sigma^2 t}\right) 2\sigma \mathrm{d}\sigma =$$

$$= \frac{2^\alpha (t\beta)^\alpha}{\sqrt{2\pi}} \frac{\Gamma\left(\frac{1}{2}+\alpha\right)}{\Gamma(\alpha)} \frac{1}{(x^2+t\beta)^{\alpha+\frac{1}{2}}} \quad . \tag{4.15}$$

We can see here the connection between volatility and temperature. This connection will be seen even more in next section.

### 4.2.3 Tsallis Entropy

When in thermodynamics we derive the most probable energy distribution for particles with given expectation value $\mathcal{E}$, we use the concept of minimal entropy. We denote $p_i$ probability that the system will be in the state $x_i \in X$, where $X$ is a set of all possible states. We demand that probability distribution should be normalized, i.e.:

$$\sum_i p_i = 1, \tag{4.16}$$

We also demand that expectation value of energy is $\mathcal{E}$, so:

$$\mathcal{E} = E(\mathcal{E}_i)_{x_i \in X} = \sum_i \mathcal{E}_i p_i, \tag{4.17}$$

where $\mathcal{E}_i$ is energy in the state $i$. The logical question is, what is the best function for entropy. The entropy function $S$ should measure the information we do not know about the system.

It should obtain these properties:

1. For $p_i = 1: S(p_i) = 0$. We know that this state is the only one that will be realized.

2. For $p_i \to 0: S(p_i) \to \infty$. When improbable state is realized, we will know much more about the system.

3. For independent states $x_i, x_j$: $S(p_i p_j) = S(p_i) + S(p_j)$. This property is sometimes called additivity of entropy.

[1] Inverse gamma distribution: continuous distribution defined on positive real line with two parameters: $\alpha$ is a shape parameter, $\beta$ is scale parameter. Probability function is :

$$p(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha) x^{\alpha+1}} \exp\left(-\frac{\beta}{x}\right).$$

The most common choice of entropy function that fulfills these properties, is $-k \ln p_i$, where $k$ is not specified constant. Because we have many particles, or events, so we take an expectation value. So we get the definition of Gibbs-Boltzmann[2] entropy:

$$S = -k \sum_i p_i \ln p_i, \tag{4.18}$$

and we try to minimize the entropy with respect to conditions (4.16), (4.17). We solve this problem by Lagrange multipliers' method. The Lagrange function is equal to:

$$\Lambda = -k \sum_i p_i \ln p_i - \alpha k \left( \sum_i p_i - 1 \right) - \beta k \left( \sum_i \mathcal{E}_i p_i - \mathcal{E} \right). \tag{4.19}$$

We find an extreme value of $\Lambda$. That means to solve system of equations:

$$\begin{aligned} \frac{\partial \Lambda}{\partial p_i} &= 0 \quad \forall i \; : \; x_i \in X \\ \sum_i p_i - 1 &= 0 \\ \sum_i \mathcal{E}_i p_i - \mathcal{E} &= 0. \end{aligned} \tag{4.20}$$

When we derive $\Lambda$, we get:

$$\frac{\partial \Lambda}{\partial p_i} = -(\ln p_i + 1) - \alpha - \beta \mathcal{E}_i = 0. \tag{4.21}$$

From this we get an expression for $p_i$:

$$p_i = \exp\left(-1 - \alpha - \beta \mathcal{E}_i\right), \tag{4.22}$$

and after normalization of $p_i$ with equation (4.16), we get well known formula for the most probable distribution:

$$p_i = \frac{1}{Z(\beta)} \exp(-\beta \mathcal{E}_i), \tag{4.23}$$

where $Z(\beta)$ is partition function and it is defined as:

$$Z(\beta) = \sum_i \exp(-\beta \mathcal{E}_i). \tag{4.24}$$

If we compare this solution with our previous remarks that $\beta \sim \frac{1}{\sigma^2}$ and energy can be representation of $(\ln S)^2$, we see that it gives us prediction that the best way, how to describe our problems will be by Wiener processes with normal distribution.

Generalization of Gibbs-Boltzmann entropy is called *Tsallis entropy* [17]. The main principle is that postulation of exactly the form of entropy shown above, gives in many cases good predictions, but there is no physical reason,

[2] also called Shannon entropy

why must logarithm be the only function suitable to describe uncertainty of the system. Tsallis gives another class of functions that can be suitable and defines entropy as:

$$S_q = -\frac{k}{1-q}\left(1 - \sum_i p_i^q\right). \tag{4.25}$$

This entropy for $q \to 1$ becomes Gibbs-Boltzmann entropy[3] . This entropy has the problem that it does not obey property 3 of suggested properties of entropy, i.e. additivity. But it is not necessarily a problem, in large open systems, as financial markets, it is possible to see that this property is not fulfilled. Here is a generalized version of additivity:

$$S_q(p_i p_j) = S_q(p_i) + S_q(p_j) + (1-q)S_q(p_i)S_q(p_j). \tag{4.26}$$

Now, we can make again Lagrange function and solve system 4.20:

$$\frac{\partial \Lambda}{\partial p_i} = -\frac{q}{q-1}p_i^{q-1} - \alpha - \alpha\beta\mathcal{E}_i = 0.^4 \tag{4.27}$$

From that we get the expression for $p_i$:

$$p_i = \alpha^{\frac{1}{q-1}}\left(\frac{1-q}{q}\right)^{\frac{1}{q-1}}(1+\beta\mathcal{E}_i)^{\frac{1}{q-1}} \tag{4.28}$$

and from normalization condition we get:

$$p_i = \frac{1}{Z_q(\beta)}(1+\beta\mathcal{E}_i)^{\frac{1}{q-1}}, \tag{4.29}$$

where partition function $Z_q(\beta)$ is equal to:

$$Z_q(\beta) = \left(\frac{1-q}{q}\right)^{\frac{2-q}{1-q}} \sum_i (1+\beta\mathcal{E}_i)^{\frac{1}{1-q}}. \tag{4.30}$$

An interesting thing about this result is the fact that this distribution is heavy-tailed. It shows that for large, open systems is more natural to use heavy-tailed distributions, although they do not have defined first or second moment.
**Note:** Another possibility is to work with *Rényi entropy* defined as:

$$S_\alpha = \frac{k}{1-\alpha}\ln\left(\sum_i p_i^\alpha\right). \tag{4.31}$$

### 4.2.4 Fractals in Financial Markets

Another point of view to the Brownian motion can be trough fractals. The founder of the fractal analysis is Benoit Mandelbrot, and he also studied applications of fractals in financial markets (e.g. [13]). Fractals are geometric objects that are self-similar under scaling. These object are usually defined recursively

[3] it can be shown by L'Hospital rule
[4] i Instead of $\beta$ we chose the second multiplyer $\alpha\beta$. That is possible for $\alpha \neq 0$.

and have special properties. Brownian motion can also be generated as fractal. If we change variables $x, t$:

$$\begin{array}{rcl} x & \mapsto & \alpha x \\ t & \mapsto & \alpha^2 t, \end{array}$$

then the probability density does not change:

$$p(\alpha x, \alpha^2 t)\ \mathrm{d}(\alpha x) = \frac{\alpha}{\sqrt{2\pi\alpha^2 t}} \exp\left(-\frac{(\alpha x)^2}{2\alpha^2 t}\right) \mathrm{d}x = p(x,t)\ \mathrm{d}x. \tag{4.32}$$

This property is called self-affinity. We could see it also from the martingale property of $W^2(t) - t$, because from that results:

$$E(W^2(t)) = t \tag{4.33}$$

All of these results lead us to the idea that Brownian motion is a fractal (all details are in [8]), and we can this motion generate as fractal.

Let us start with a line shown in the first picture of 4.5. It is called initiator. The line symbolizes a drift of an asset price given by parameter. In every iteration we divide the line into three parts. These lines are called generator. The first part will be from 0 to $\frac{4}{9}$ of the length of the line, the second will be from $\frac{4}{9}$ to $\frac{5}{9}$ and the last part will be the rest of the line. Now we replace the line with three lines. We denote the boundary coordinates of the line $(0,0)$,$(d,d)$, then the first line be given by points $(0,0)$, $(\frac{4}{9}d, \frac{2}{3}d)$, the second line will be given by: $(\frac{4}{9}d, \frac{2}{3}d)$ ,$(\frac{5}{9}d, \frac{1}{3}d)$ and the last line will have coordinates $(\frac{4}{9}d, \frac{2}{3}d)$, $(d,d)$. In next iterations every line breaks to three lines in the same way. In 4.5 are shown first three iterations of this pseudo-Brownian motion.

Note that if we look at any line and compare differences of x-coordinates and y-coordinates, then the ratio will be:

$$\frac{\Delta y}{\Delta x} = \frac{l}{l^2}, \tag{4.34}$$

which corresponds to the properties of Brownian motion, where $\Delta x \sim \Delta t$ and $\Delta x \sim \Delta(\ln S)$.

Instead of fractals it is possible to use multifractals. The multifractal is a geometric object that consists more fractal structures[5]. That means, we can choose a few generators, and in every step use another generator. That seems to be effective way of modeling more realistic processes.

The formalism of multifractals is connected with the concept that says that market time passes differently from clock time. We can see that in some time periods, especially when a stock opens, or just before closing, we can see much more trades, so the prices of financial products change their values much more in these time periods. So we can imagine that in the morning and before closing passes trading time slower, than clock time, and in the lunch time when many traders are not available, passes trading time faster.

When we know, how behave markets during the day, we can use an appropriate generator. This concept can generate very realistic behavior of asset prices [13].

[5]This is not the definition of multifractal, for more details, see [8].

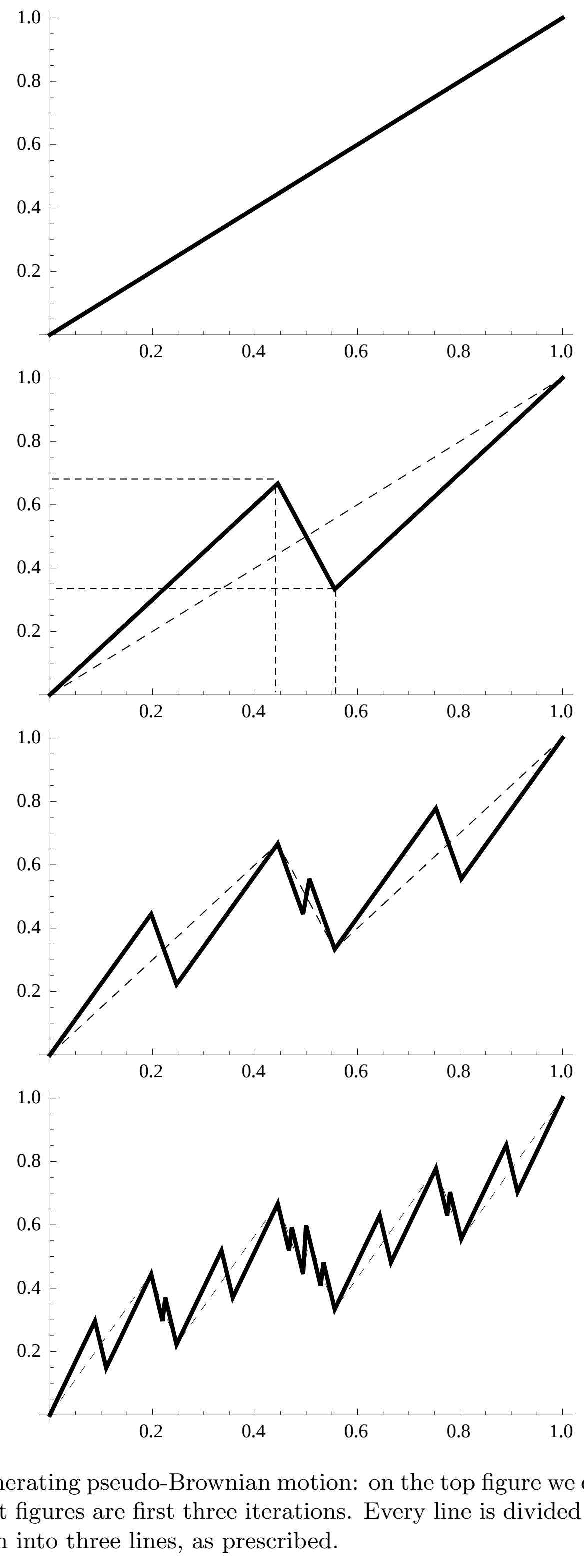

Figure 4.5: Generating pseudo-Brownian motion: on the top figure we can see initiator, and on the next figures are first three iterations. Every line is divided into three parts and it is broken into three lines, as prescribed.

# Bibliography


[1] Louis Bachelier. *Théorie de la spéculation.* Gauthier-Villars, 1900.

[2] Martin Baxter and Andrew Rennie. *Financial Calculus: An introduction to derivative pricing.* CUP, Cambridge, 2001.

[3] Christian Beck. Recent developments in superstatistics. *Brazilian Journal of Physics*, 39(2A), 2009.

[4] Christian Beck and E. D. G. Cohen. Superstatistics. *Physica A*, (322), 2003.

[5] Jean-Phillipe Bouchard and Marc Potters. *Theory of Financial Risks: From Statistical Physics to Risk Management.* CUP, Cambridge, 2001.

[6] Zdzisław Brzeźniak and Tomasz Zastawniak. *Basic Stochastic Processes.* Springer, London, 1999.

[7] Richard Durret. *Stochastic Calculus: A Practical Introduction.* CRC Press, Inc., 1996.

[8] Kenneth Falconer. *Fractal Geometry: Mathematical Foundations and Applications.* Wiley, Inc., 1990.

[9] Willian Feller. *An Introduction to Probability theory and its Applications*, volume 1. Wiley, 3. edition, 1967.

[10] Willian Feller. *An Introduction to Probability theory and its Applications*, volume 2. Wiley, 2. edition, 1970.

[11] B. V. Gnedenko and A. N. Kolmogorov. *Limit Distributions for Sums of Intependent Random Variables.* Adison-Wesley, 1968.

[12] Steven L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2), 1993.

[13] Benoit Mandelbrot. Fractal financial fluctuations; do the threaten sustainability? In *Science for Survival and Sustainable Development.* Pontificia Academia Scientiarum, 1999.

[14] Rosario N. Mantegna and H. Eugene Stanley. *An Introduction to Econophysics.* CUP, Cambridge, 2000.

[15] Leonard Ornstein and George Eugene Uhlenbeck. On the theory of the brownian motion. *Physical Review*, 36, 1930.

[16] Wolfgang Paul and Jörg Baschangel. *Stochastic Processes: From Physics to Finance*. Springer, Berlin, 1999.

[17] Constantino Tsallis, Silvio Lévy, André Souza, and Roger Maynard. Statistical-mechanical foundation of the ubiquity of lévy distributions in nature. *Physical Review*, 75(20), 1995.