January 14, 2002 (Preliminary Version 1.0)

The Changing Concept of Financial Risk

Abstract

The recent rapid accumulation of anomalous empirical research results has made clear that the classical definition of financial risk based on asset classes only is ready for a epistemological change. Currently, the definition of financial risk suffers from three major deficiencies: (1) financial risk is insufficiently measured by the conventional second - order moments; (2) financial risk is assumed to be stable and all distribution moments are assumed to be time-invariant;.and (3) pricing observations are assumed to exhibit only serial dependencies, which can be simply removed by simple inverse transformations, like the geometric Brownian motion, Markov, ARIMA, or (G)ARCH models. But (1) higher - order moments are acknowledged by experienced traders to be influential in asset and derivative valuations; (2) distributions of returns are observed to be nonstationary; and (3) dificult to observe long term dependencies have surprised many risk managers. Based on accumulated empirical evidence, a new funtional definition of financial risk that takes account of asset classes, time and time horizons is required to fully capture the concept as required in the empirical financial markets.

Cornelis A. Los
Kent State University, College of Business Administration and Graduate School of Management
BSA 416, Kent, OH 44272, USA
Email: clos@kent.edu

1 INTRODUCTION

1.1 Classical Market Returns Assumptions

Most investors, portfolio managers, corporate financial analysts, investment bankers, commercial bank loan officers, security analysts and bond-rating agencies are concerned about the uncertainty of the returns on their fundamental investment assets, caused by the variability in speculative market prices (market risk) and the instability of business performance (credit risk).[1]

Derivative instruments have made hedging of such risks possible. Hedging allows the selling of such risks by the *hedgers*, or suppliers of risk, to the *speculators*, or buyers of risk, but only when such risks are predictable, *i.e.*, when they show a certain form of inertia or stability. Indeed, the current derivative markets are regular markets where "stable," *i.e.*, predictable risk is bought and sold.

Unfortunately, all these financial markets suffer from three major deficiencies:

1. Risk is insufficiently measured by the conventional second - order moments (variances and standard deviations). Often one thinks it to be sufficient to measure risk by only second - order moments, because of the facile, but erroneous, assumption of normality (or Gaussianness) of the price distributions produced by the market processes of shifting demand and supply curves.

2. Risk is assumed to be stable and all distribution moments are assumed to be invariant.

3. Pricing observations are assumed to exhibit only serial dependencies, which can be simply removed by appropriate transformations, like the geometric Brownian motion, Markov, ARIMA, or (G)ARCH models.

Based on these simplifying assumptions, investment analysis and portfolio theory have con-

[1] With little loss of generality, we mean by asset or investment returns: total returns = sum of cash payments and capital gains. All dividend payments are assumed to be reinvested in the asset.

ventionally described financial market risk as a function of asset class only. In a simplifying representation:

$$\boxed{\textbf{portfolio return volatility } \sigma_{pp} = f(\text{asset class})} \tag{1}$$

Fig. 1 shows this classical presentation of financial risk as a function of asset class by Ibbotson and Sinquefield, who have collected annual rates of return as far back as 1926 (Ibbotson and Sinquefield, 1999). The dispersion of the return distributions, measured by the respective standard deviations, differs by six different asset classes:

(1) common stocks of large companies

(2) common stocks of small firms

(3) long - term corporate bonds

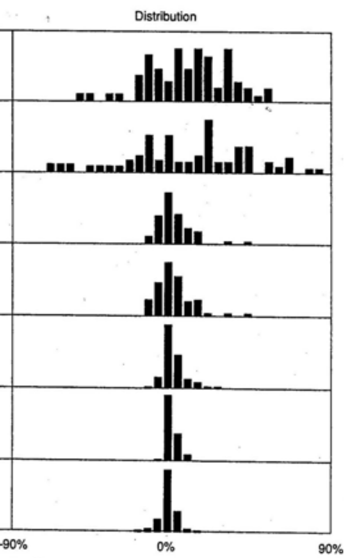
(4) long - term U.S. government bonds

(5) intermediate - term U.S. government bonds

(6) U.S. Treasury bills

When an investor wants a higher return combined with more risk, he invests in small stocks, when he wants less risk and accepts a lower return, he is advised to invest in cash.

	Geometric Mean	Arithmetic Mean	Standard Deviation
Large company stocks	10.2%	12.2%	20.3%
Small company stocks	12.2	17.4	34.6
Long-term corporate bonds	5.4	5.7	8.4
Long-term government bonds	4.8	5.2	8.8
Intermediate-term government bonds	5.1	5.2	5.7
U.S. Treasury bills	3.7	3.7	3.3
Inflation	3.1	3.2	4.6



*The 1933 Small Company Stock Total Return was 142.9 percent.

Source: © *Stocks, Bonds, Bills, and Inflation 1996 Yearbook,*™ Ibbotson Associates, Chicago (annually updates work by Roger G. Ibbotson and Rex A. Sinquefield). Used with permission. All rights reserved.

Historical Average Annual Returns and Return Variability, 1926 - 1995

For example, Tobin's (1958) *liquidity preference theory* shows that any investment risk level (as defined by the second moment of asset returns) can be attained by a linear combination of the market portfolio and cash, combined with the ability to hold short (borrow) and to hold long (invest). The market portfolio contains all the non - diversifiable systematic risk, while the cash represents the "risk - free" asset, of which the return compensates for depreciation of value caused by inflation. The linear combination of the market portfolio and cash can create any average return and any risk - premium one wants or needs, under the assumption that the distributions of these investment returns are mutually independent over time.

1.2 What's Empirically Wrong?

Regrettably, there are many things wrong with this oversimplified conceptualization and modeling of financial market risk and one has now become alarmingly obvious. For example, financial disasters are much more common and occur with much higher frequencies than they should be according to the classical assumptions. An incomplete but rather convincing listing of financial disasters can be found in Kindleberger (1996). *Cf.* also Bernstein (1996), Schroeder (1991) and Bassi, Embrechts and Kefetzaki (1998). The world's financial markets exhibit longer term pricing dependencies, which show, like the Plagues of the Old Testament, devastating, but essentially unpredictable non - periodic cyclicity, or sharp and disastrous discontinuities, like catastrophic floods, in aggregated trading observations, or turbulence structures and "eddie" like condensation and rarefaction patterns in high frequency trading frequencies.

First, we'll quickly learn that the uncertainty of the investment returns is a much wider concept than just the volatility of the prices as measured by second - order moments. Higher - order moments, like *skewness* and *kurtosis*, play a still undervalued, but very important role. For example, the distributions of investment returns exhibit positive biases, because of the termination of non - performing businesses and the continuing life of performing ones. In addition, the tails of the rate of return distribution returns are fatter, *i.e.*, the outlying returns are more prevalent, than normally expected.

Second, we will observe that stationarity of the investment returns cannot be so easily assumed, since we empirically observed that the distributions of investment returns change over time. This probably occurs because markets develop their institutional frameworks and mature, thereby changing the constraints of their financial pricing processes.

Third, we find that intertemporal dependencies cannot easily be filtered out of the observed pricing series. The random pricing processes cannot be so easily reduced to independent white noise series, since financial pricing series exhibit global dependencies. In fact, Wald's (1943) assumption of *serial* dependence with short term memory can be shown to be empirically false. Global, long term dependence plays a pervasive and important role. Thus, instead, it is better to present financial market risk, in a simplifying representation, as follows:

$$\boxed{\textbf{asset return distribution} = f(\text{asset class, time horizon } \tau,\, t)} \tag{2}$$

Thus, not only is the price distribution produced by the speculative markets dependent on the asset classes and on the time horizons τ of the investors, but this distribution function may be time-varying, as indicated by the time $t-$argument. This empirical reality, which only recently is becoming properly modeled, has serious consequences for portfolio management and investment analysis, since Tobin's (1958) liquidity preference theory is clearly too simple to adequately reflect all these dimensions of risk. The simple, static, two - dimensional return - risk trade - off, on which classical Modern Portfolio Theory (MPT) is based, will have to be replaced by multi - dimensional and dynamic return - risk trade - offs, as was earlier suggested in Los (1998, 2000b).

Example 1 *An example of the time - dependence of price distributions is the strong time - dependence of the standard deviation, or volatility, of stock price changes (Schwert, 1989).*

This paper contains many definitions to acquire a proper analytic and technical lingo to discuss the necessity of a new definition of financial risk. In particular, we'll review Kolmogorov's classical axiomatic (set-theoretic) definition of probability and of random processes and contrast them with the much broader definitions of uncertainty and pure randomness. We also look at empirical definitions of frequency distributions and of observed time series, and at the summarizing characterizations of these time series by their often time-varying moments and cumulants.

2 UNCERTAINTY

There is no doubt in the mind of physicists that uncertainty, like relativity, is of an absolutely fundamental nature, that admits no exceptions. The world could not even physically exist without uncertainty:

> "One of the fundamental consequences of uncertainty is the very size of atoms, which, without it, would collapse to an infinitesimal point." (Schroeder, 1991, p. 113)

In mathematics, the theory of Hilbert bases and (linear) operator algebra led to the formulation of the *Uncertainty Principle* (Meyer, 1987), which I judiciously and fruitfully exploited in my own book (Los, 2001). But for the future development of financial risk theory we need a broader definition of uncertainty.

According to Webster's *New Universal Unabridged Dictionary* (Deluxe Second Edition), Dorset & Baber, 1983, p. 1990):

un·cĕr 'tain·ty = the quality or state of being uncertain; lack of certainty; doubt

and

un · cĕr 'tain = **1.** not certainly known; questionable; problematical.

2. vague; not definite or determined.

3. doubtful; not having certain knowledge; not sure.

4. ambiguous.

5. not steady or constant; varying.

6. liable to change or vary; not dependable or reliable

Similarly, in modern risk theory, we distinguish three different, but closely related concepts: randomness, chaos and probability.[2] Let's explain what each of these concepts mean and discuss their limitations.

2.1 Randomness = Irregularity

Essentially, from Webster's Dictionary, we have the following informal definition for randomness[3]

---

[2] Surprisingly, these crucial distinctions were already made in 1921 in a non-mathematical form by two economists, Frank Knight (1885 - 1972) and John Maynard Keynes (1883 - 1946). In particular, the Chicago economist Frank Knight made sharp distinctions between uncertainty, randomness, and probability. The commonsensical Knight would have appreciated the modern concept of randomness as irregularity, which we adopt as the most rational, although, perhaps, not the exact measurement of it. The rather elitist Cambridge, UK economist Keynes distinguished between historical probability = relative frequency, and subjective probability. Keynes personally appreciated the subjective probability concepts developed by the 18th century enigmatic, nonconformist minister Thomas Bayes (1701 - 1761), a Fellow of the Royal Society, in Bayes ' posthumously published "Essay Towards Solving A Problem In The Doctrine of Chances " in *Philosophical Transactions* (1763). See also section 4.1.3, "Conditional Probability."

[3] John Stuart Mill (1806 - 1873), the great 19th century English philosopher and economist demanded in his influential pamphlet *On Liberty (1859)*, in which he relentlessly attacked conformity and timidity, that we accept

randomness = the state of being haphazard, not unique, or irregular

Clearly this simple definition is based on the common use of the word "randomness" in the English language (*Cf.* Bernstein, 1996), but it does not provide us with much of a basis for the measurement of randomness. For example, how do we distinguish between genuine and pseudo-randomness? Cryptographers are very much interested in this distinction.

2.1.1 Degree of Irregularity

What is regular is defined, fixed and clearly determined. But how irregular is the absence of determinedness? Recently, Pincus and Singer (1996) asked the question: what is the degree of irregularity in time series and how do we measure that? One extreme is the certainty of being fixed, of being unique, a constant, and having thus no spectrum at all. The other extreme, the ultimate state of irregularity, is when something is indistinguishable from background noise, that has no spectral features, *i.e.*, noise that covers the whole spectrum. For example, white noise has a *flat* spectrum. Thus, it exhibits a specific, distinguished spectral feature and, therefore, cannot be called irregular or random background noise. In between the two extremes we find degrees of irregularity which can be described by variously shaped spectra. Each irregular series has its own spectrum, be it a Fourier spectrum for stationary series; or a changing spectrum for nonstationary series, to be analyzed by either short-term Fourier Transforms or by Wavelet Transforms, depending on how fast the changes occur; or a so-called singularity spectrum for observational series which show many discontinuities or jumps.

2.1.2 Measures of Sequential Irregularity

In financial risk theory we are not interested in irregularity per se, but in dynamic irregularity, *i.e.*, in irregularity as it manifests itself over time. For example, how irregular are the prices produced

uncertainty or randomness as the necessary human condition. He wanted us to live, as his modern disciple Isaiah Berlin might say, with the assumption that life is neither stationary nor easily understood. In this context, it is significant that in his first major work, *A System of Logic* (1943), Mill analyzed the epistemological principles underlying empiricism.

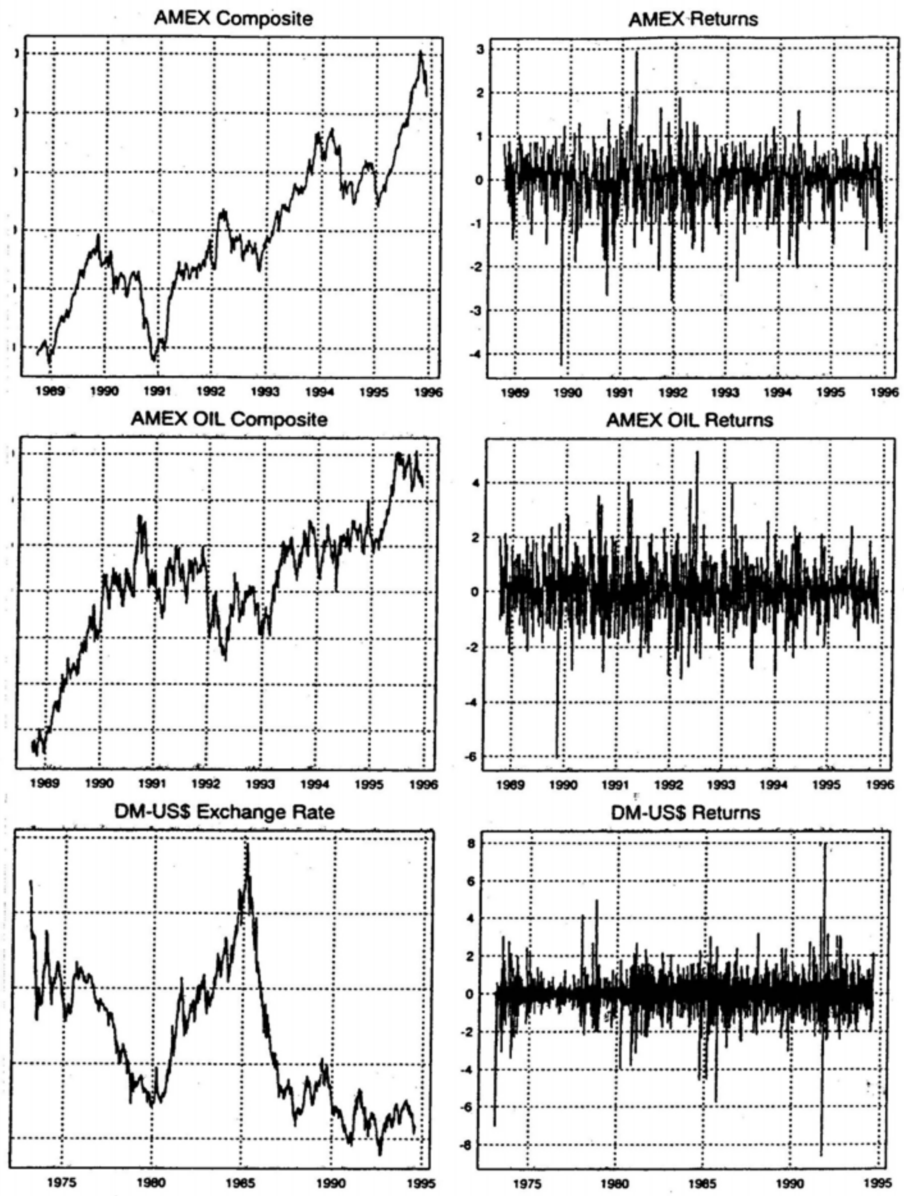
by a market pricing mechanism over time? Fig. 2 provides some examples of financial market price series and their rates of return:[4]

1. the daily AMEX Composite index from September 1, 1988 to July 28, 1994, with $T = 1,810$ observations;

2. the daily AMEX OIL Composite index from September 1, 1988 to July 28, 1994, with $T = 1,810$ observations; and

3. the daily DM/USD Exchange Rate from January 2, 1973 to July 28, 1994 with $T = 5,401$ observations.

[4] Source: Mittnik, Rachev and Paolella (1998), p. 84.



Levels and returns of empirical financial time series

For the past hundred years, since Bachelier's Ph.D. thesis of 1900, in which he described speculative price formation as a random walk, people have attempted to describe the degree of irregularity of market pricing and of related investment returns. The currently best known rational measures of such irregularity are the Lipschitz exponents (such as the Hölder and Hurst exponents).

2.2 Pseudorandomness versus Genuine Randomness

People working with computers often sloppily talk about their system's "random number generator" and the "random numbers" it produces. But numbers calculated by a computer through a deterministic process, cannot, by definition, be random. Given knowledge of the algorithm used to create the numbers and its internal state, you can predict all the numbers returned by subsequent calls to the algorithm, whereas with genuinely random numbers, knowledge of one number or of an arbitrarily long sequence of numbers is of no use whatsoever in predicting the next number to be generated. In the first case a key will open all random numbers for you, in the second case no key will.

Computer - generated "random" numbers are more properly referred to as *pseudorandom* numbers, and pseudorandom sequences of such numbers. A variety of clever algorithms have been developed which generate sequences of numbers which pass every current statistical test used to distinguish random sequences from those containing some pattern or internal order. A high-quality pseudorandom sequence generator generates data that are indistinguishable from a sequence of bytes chosen at random. Indistinguishable, but not genuinely random!

We no longer have to use pseudorandom number generators to simulate "randomness." There are now systems to collect *genuine random numbers*, generated by a process fundamentally governed by the inherent uncertainty in the quantum mechanical laws of nature, directly to your computer in a variety of forms.

Example 2 ***Hotbits*** *are random numbers generated by timing successive pairs of radioactive decay electrons or beta particles. These particles are produced by the spontaneous transformation of neutrons (with charge* 0*) in the nucleus of Krypton*-85 *into pairs of protons (with charge* $+1$*) and free electrons (= beta particles with charge* -1*). The free electrons, or "beta rays," are then detected by a Geiger - Müller tube in a simple radiation monitor (Fig. 3) interfaced to a computer. The unstable nucleus of the radioactive Krypton*-85 *(the* 85 *means there are a total of* 85 *protons and neutrons in the atom) spontaneously turns into the stable nucleus of the non-radioactive Rubidium*-85 *which still has a sum of* 85 *protons and neutrons, and a beta particle is emitted with an energy of* 687 *kiloelectron volts (keV), resulting in no net difference in charge:*

$$^{85}Kr \longrightarrow ^{85}Rb + \beta^{-} + \gamma \tag{3}$$

In this case, a gamma ray is also emitted with an energy of 514 *keV, carrying away some of the energy. "Gamma rays" are nothing other than photons – particles of light, just carrying a lot more energy than visible light. Krypton*−85 *has a half life of* 10.73 *years. This is called its half - life, since every* 10.73 *years half of a very large number of Krypton*−85 *nuclei present at the start of the period have decayed into Rubidium*−85*. But there is no way, even in principle, to predict when a given atom of Krypton*−85 *will decay into Rubidium*−85*. It has a* 50/50 *chance of doing so in the next* 10.73 *years, but that's all we can say. The inherent uncertainty of such decay time is genuinely random. Since the time of any given decay is random, the interval between two consecutive decays is also genuinely random (not unlike between two financial transactions). Using the Geiger teller, we can now measure the lengths of the uncertain intervals after the fact and thus collect genuinly random numbers.*

Simple radiation monitor

We measure a pair of these intervals, and emit a zero or one bit based on the relative length of the two intervals. To create each random bit, we wait until the first count occurs, then measure the time, T_1, until the next. We then wait for a third pulse and measure T_2, yielding a pair of durations. If they're the same, we throw away the measurement and try again. Otherwise if T_1 is less than T_2 we emit a zero bit; if T_1 is greater than T_2, a one bit. In practice, to avoid any residual bias resulting from non - random systematic errors in the apparatus or measuring process consistently favouring one state, the sense of the comparison between T_1 and T_2 is reversed for consecutive bits.

2.3 Chaos = Deterministic Dynamic Non-uniqueness

Chaos is a special form of irregularity. It means that at a certain time something, that was certain and unique, suddenly can become non - unique, although it remains very well determined. The system can have more than one equilibrium state to be in, for example, because of equilibrium state *bifurcations*. How many times it "orbits" or "jumps" through such a set of equilibrium states depends on the nonlinear constraints imposed on the dynamic system. However, these separate equilibrium states of the same system can be perfectly well identified, determined and described, like H_2O molecules orbiting through the two coexisting equilibrium states of ice floating in liquid water. The molecule is either in the "ice" state, or in the "water" state, and it orbits through these two states as time progresses. Thus

$$\boxed{\textbf{chaos} = \text{deterministic dynamic non-uniqueness}}$$

Chaos is a form of randomness, since there is clearly non - uniqueness and non - periodicity. But it is not pseudo - randomness, since there is more than one state of being at the same time, while with pseudorandomness there is only one state of being at the same time. However, chaos is deterministic, since there is no doubt what these distinct but coexisting equilibrium states are: they are perfectly well determined. The set of such distinct but coexisting non - periodic equilibrium states is called the *strange attractor*. Already more than a decade ago, such chaos has been observed in speculative market prices on the trading floor (Savit, 1988, 1994).

2.4 Probability = Complete Set of Relative Frequencies

Probability is a very well defined, complete and constraint form of "randomness". In mathematics, probability is the ratio of the chances favoring a certain state to all the chances for and against it. Thus, probability is a *rational measure:* it measures *relative frequencies.* It counts the number of times of being in state A relative to the total number of states, *i.e.*, the sum of the number of times of being in state A relative to the number of times of being in the $non - A$ states.

Remark 3 *Earlier financial analysts claimed that probability is a necessary and sufficient concept for the pricing of derivatives and that the broader concept of uncertainty would not be necessary or sufficient. Nowadays it is acknowledged that probability is not needed for the pricing of options, because the prices of derivatives can be replicated by the prices of portfolios of fundamental assets, and the prices of fundamental assets are uncertain. Even bond prices are uncertain, when discontinuous credit or default risk is taken into account. However, such uncertainty does not prevent one from giving an ex post* ***probabilistic interpretation*** *to the pricing of derivatives.*

2.4.1 Kolmogorov's Axiomatic Probability

In 1933 the Russian mathematician Kolmogorov provided one possible, formal axiomatic definition of probability, using set theory (Kolmogorov, 1933; *cf.* also Papoulis, 1984, for a complete treatment). Since Kolmogorov, particular non - Kolmogorovian definitions of probability have been discovered, *i.a.*, by the Italian mathematician Luigi Accardi, a student of Kolmogorov and now Professor of Mathematics at the University of Rome. Thus, currently, there co - exist several definitions of probability. The definition of probability is no longer unique! However, the financial literature still presents only the simple definitions of Kolmogorov's set theoretic definition of probability, since they are the most familiar, although somewhat misleading representation of "randomness," and they assist with making the connection to the modern approach to randomness, *i.e.*, to our broader concept of randomness as "irregularity."

To provide the axiomatic definition of probability in the proper context of set definitions, one needs some familiarity with the Boolean notions of *sets*, *complements*, *intersections* and *unions*, as well as with the notion of *function*. Usually Kolmogorov's definition is gradually arrived at by a sequence of definitions sequentially building on each other, beginning with the fundamental set definitions (Chow & Teicher,1978; Los, 1982). One needs, somewhat paradoxically, the definition of *almost certainty* in order to define Kolmogorov's very restricted concept of randomness.

Definition 4 *A property is said to hold* ***almost certainly*** *(a.c.) [also: almost surely (a.s.), almost everywhere (a.e.), or with probability one (w.p.1)], if it holds everywhere, except possibly on a null set A, i.e., a set A such that* $P(A) = 0$.

Kolmogorov's, axiomatic definition of a random variable is as follows.

Definition 5 *(Kolmogorov) Let* (Ω, G, P) *be a complete probability space. A measurable function* $X\colon \Omega \to \mathbb{R}$ *(where* $\mathbb{R}$ *is the real line) is said to be a* ***random variable*** *(r.v.) if*

$$P(A) = 0, A = \{t : |X(t)| = \infty\} \tag{4}$$

Surprisingly, perhaps, Kolmogorov's apparently innocuous definition excludes a lot of irregular events, *e.g.*, all singularities and discontinuities. Thus, nowadays, Kolmogorov's definition is considered deficient by real world mathematicians, such as Mandelbrot, Pincus, Kalman and Singer.

For example, we can observe and measure (= count the frequency of) special irregularities, called "singularities." Nowadays, mathematicians even discuss the concept of a "singularity spectrum" and how to measure it. Thus this new non - Kolmogorov concept of "probability" effectively asserts:

$$P(A) \geq 0, A = \{t : |X(t)| = \infty\} \tag{5}$$

2.4.2 Real World: Relative Frequency

In the real, empirical world, we determine the probability of an irregular event by measuring how relatively often it occurs. We measure its relative frequency of occurrence. We do not require that the absolute measure of that event is finite. Thus, for empirical simplicity, we have a measurable concept of definition of probability as:

probability = relative frequency of events

This measurable definition focuses on the weakness of Kolmogorov's definition of probability: you have to know the universe, the complete basis for your relative frequency. What is that universe of possible events and how do we know it is complete?

Based on the preceding discussion of why probability measures a very specialized, "complete" form of randomness, we categorically state that it is more in agreement with empirical reality to define

randomness = *irregularity*

and not to bound the describing and measuring functions of such irregularity, in contrast to what happens in Kolmogorov's probability theory. Irregular is everything that is not regular. Polynomial functions, like higher-order functions of time, are considered regular and systematic.

Any financial phenomenon that can not be explained or known in the form of such a polynomial is then considered irregular or "risky."

2.5 The Ellsberg Paradox

Peters (1999) uses the Ellsberg Paradox to illustrate the essential difference between uncertainty and probabilistic risk.[5] In the Ellsberg Paradox, you are shown an urn that contains 90 balls. Of these, 30 balls are red, and the remaining 60 balls are an *unknown* mixture of white and blue balls. One ball is to be drawn from the urn, and you are paid an amount of money if a particular color ball is chosen. You are given two payoff options to choose from as in Table 2.

TABLE 1: ELLSBERG PARADOX PAY-OFFS A

	Red	**White**	**Blue**
Option 1	$100	$0	$0
Option 2	$0	$100	$0

Look over Options 1 and 2 in Table 1 and decide which you would choose and keep your choice in mind. Most people choose Option 1 for the set of payoffs in Table 1. We will soon see why.

Next, turn to two other options are offered in Table 2. The drawing will be of the same urn, with the same mixture of red, white and blue balls, as before.

TABLE 2: ELLSBERG PARADOX PAY-OFFS B

	Red	**White**	**Blue**
Option 3	$100	$0	$100
Option 4	$0	$100	$100

Which of these two new options would you now choose? Be honest! Most people choose Option 4 for the second set in Table 2. But why?

In Option 1, you know for certain that you have a $\frac{1}{3}$ probability of winning. But you have no idea of the probability of winning in Option 2. It could be anywhere between zero and $\frac{2}{3}$, i.e., it

[5] This section is borrowed, with slight alterations, from Edgar Peters (1999, pp. 22 - 24). Dr. Daniel Ellsberg is, indeed, the one of Pentagon Papers' fame.

could be zero or higher than the $\frac{1}{3}$ of Red. Which demonstrates that most people prefer to go with the odds you know, instead of choosing for uncertainty.

Option 4 is chosen for the same reason. You know that Option 4 has a $\frac{2}{3}$ probability, because 60 of the 90 balls are either white or blue, but you do not know the odds of finding a red *or* a blue ball, which can be anywhere between $\frac{1}{3}$ and 1. Again, most people prefer to go with the odds you know, instead of facing uncertainty.

Now, by itself each choice appears rational. Remember, though, that you chose *both* Options 1 and 4. Here is where the Paradox comes into play. Choosing Option 1 over Option 2 means that you *believe* that a red ball is more likely to be drawn than a white ball. However, choosing Option 4 over Option 3 implies that you *believe* that "white or blue " is more likely than "red or blue " and thus you *believe* that white is more likely than red when choosing Option 4. Consequently, choosing both Option 4 and Option 1 under the same conditions is irrational, since the two beliefs supporting the respective choices are in conflict with each other, according the tenets of subjective probability, so beloved by "rational" Bayesian statisticians.

Why do the majority of people choose Options 1 and 4? Because, when faced with the unknown, or true uncertainty, we are more comfortable with what we know, than what we don't know. Thus, uncertainty is very different from probabilistic risk, which is known. Probabilistic risk depends on the concept of known odds. The odds are known and calculable for probabilistic games like throwing dice, turning a wheel of fortune, or playing a hand of cards. But uncertainty is more dangerous than low - but - known odds, since, essentially,

uncertainty = our ignorance

We hate being ignorant! It is this hate for ignorance that drives scientists to look for certain and unique mathematical models to explain the structure of natural phenomena. Once such a mathematical model is found, *e.g.*, DNA's double helix, there is no longer uncertainty (Cf. Los, 2001, Chapter 1).

3 NON-PARAMETRIC AND PARAMETRIC DISTRIBUTIONS

Let's introduce a few additional classical measurement definitions to underpin the conventional the frequency distribution analysis. We define first the classical time - independent distribution function.

Definition 6 *Let X be a random variable defined on the probability space (Ω, G, P). Then the **distribution function** (d.f.) of X is defined by the probability (= relative frequency) P such that*

$$F_X(x) = P\{t : X(t) < x, x \in [-\infty, \infty)\} \tag{6}$$

Remark 7 *From the preceding discussion it is clear that the distribution function (d.f.) has the following properties:*

$$\text{(i) } F_X \text{ is nondecreasing} \tag{7}$$

$$\text{(ii) } F_X \text{ is left continuous: } \lim_{\substack{y<x \\ y \to x}} F(y) = F(x),\ x \in R \tag{8}$$

$$\text{(iii) } F_X(\infty) = \lim_{x\to\infty} F_X(x) = 1$$

$$F_X(-\infty) = \lim_{x\to\infty} F_X(-x) = 0 \tag{9}$$

This distribution function could easily be made a function of time t and of time horizon (or time lag) τ by writing $F_X(t, \tau)$. By itself, such a definition of a frequency distribution is a sufficiently flexible concept, which allows the extension of the definition of financial risk. Signal procesing engineers, using Kalman filtering, have used this extended definition already for more than fourty years.

3.1 Moment and Cumulant Generation

The complete description of empirical distributions requires infinite knowledge, which we, humans, don't possess. Therefore, we try to summarize such distributions by using a limited number of characterizing *summary statistics*. We will discuss the definitions and properties of moments and cumulants, in particular, of the first four orders of distributions. These moments are useful semi - invariant summary statistics of distributions.

Definition 8 *Given a set of n real random variables* $\{x(1), x(2), ..., x(n)\}$, *their* ***joint moments*** *of order* $r = k_1 + k_2 + ... + k_n$ *are given by the partial derivatives of the characteristic function evaluated at zero frequencies* $\omega_i, i = 1, 2, ..., n$ *(Papoulis, 1984):*

$$\begin{aligned} Mom\left[x^{k_1}(1), x^{k_2}(2), ..., x^{k_n}(n)\right] &= E\left[x^{k_1}(1).x^{k_2}(2)...x^{k_n}(n)\right] \\ &= (-j)^r \frac{\partial^r \Phi(\omega_1, \omega_2, ...\omega_n)}{\partial\omega_1^{k_1}\partial\omega_2^{k_2}...\partial\omega_n^{k_n}} \mid_{\omega_1=\omega_2=...\omega_n=0} \end{aligned} \tag{10}$$

where

$$\Phi(\omega_1, \omega_2, ...\omega_n) = E\left\{e^{j(\omega_1 x(1)+\omega_2 x(2)+...+\omega_n x(n))}\right\} \tag{11}$$

is their ***joint characteristic function***, $E\{.\}$ *denotes the expectation operation, the number* $e = \lim_{m\to\infty}(1+\frac{1}{m})^m = 2.71828....$, *and* j *is the imaginary number* $j = \sqrt{-1}$ *(or* $j^2 = -1$*).*

Definition 9 *For one continuous random variable* X *with density* $f(x)$, *the* ***characteristic function*** *is the Fourier Transform of the density function* $f(x)$ *defined by*

$$\Phi(\omega) = E\left\{e^{j\omega x}\right\} = \int_{-\infty}^{+\infty} e^{j\omega x} f(x)dx \tag{12}$$

The characteristic function completely determines the distribution of X *and has many useful mathematical properties.*

For example, for two joint random variables $\{x(1), x(2)\}$, we have the second - order moments

$$Mom\left[x(1), x(2)\right] = E\left\{x(1).x(2)\right\}, \tag{13}$$

$$Mom\left[x^2(1)\right] = E\left\{x^2(1)\right\} \tag{14}$$

$$\text{and } Mom\left[x^2(2)\right] = E\left\{x^2(2)\right\} \tag{15}$$

However, often in signal processing, instead of using moments, cumulants are used, because of their ability to suppress noise, when it is additive Gaussian, and their usefulness for estimating frequencies. Let's first define these cumulants.

Definition 10 *The* ***joint cumulants*** *of order* r *of the same set of random variables are defined as the coefficients in the Taylor expansion of the natural logarithm of the characteristic function about zero, i.e.,*

$$Cum\left[x^{k_1}(1), x^{k_2}(2), ..., x^{k_n}(n)\right] = (-j)^r \frac{\partial^r \ln\left[\Phi(\omega_1, \omega_2, ...\omega_n)\right]}{\partial\omega_1^{k_1}\partial\omega_2^{k_2}...\partial\omega_n^{k_n}} \mid_{\omega_1=\omega_2=...\omega_n=0} \tag{16}$$

In this paper, I will not discuss multi - variate joint distributions of the set of random variables $\{x(1), x(2), ..., x(n)\}$, but only distributions of a single random variable x, a rate of return on an asset. In that case their are no cross moments, and the moments m_r can be simply computed by

$$m_r = E\left\{x^r\right\} = \int_{-\infty}^{+\infty} x^r f(x)dx \tag{17}$$

where $f(x)$ is the *probability density function* (p.d.f).

For example, the first four moments of x are simply

$$m_1 = Mom\,[x] = E\,\{x\} \tag{18}$$

$$m_2 = Mom\,[x.x] = E\,\{x^2\} \tag{19}$$

$$m_3 = Mom\,[x.x.x] = E\,\{x^3\} \tag{20}$$

$$m_4 = Mom\,[x.x.x.x] = E\,\{x^4\} \tag{21}$$

and the first four cumulants of x are related to these moments, as follows:

$$c_1 = Cum\,[x] = E\,\{x\} = m_1 \tag{22}$$

$$c_2 = Cum\,[x.x] = m_2 - m_1^2 \tag{23}$$

$$c_3 = Cum\,[x.x.x] = m_3 - 3m_2m_1 + 2m_1^3 \tag{24}$$

$$c_4 = Cum\,[x.x.x.x] = m_4 - 4m_3m_1 - 3m_2^2 + 12m_2m_1^2 - 6m_1^4 \tag{25}$$

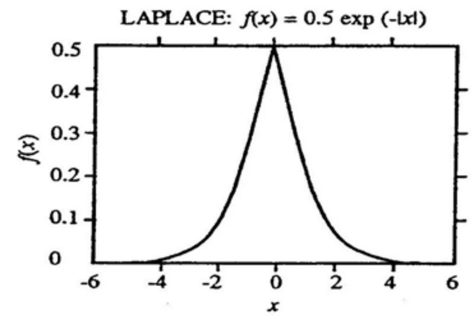
These relationships between the moments and cumulants can be verified by substituting the Taylor expansion into the preceding general definitions for joint moments and cumulants and working out the differentiations about zero. Notice that if the first moment (mean) $m_1 = E\,\{x\} = c_1 = 0$, there is considerable simplifications, since it follows that $c_2 = m_2$, $c_3 = m_3$, and $c_4 = m_4 - 3m_2^2$.[6]

3.1.1 Moments of Parametric Distributions

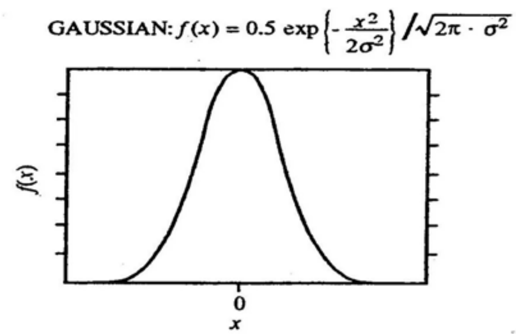
The following Figs. 4 and 5 are from Nikias and Petropulu (1993), pp. 10 and 11. Fig. 4 illustrates the first four order moments and cumulants of the p.d.f.s for three symmetric parametric distributions: the *Laplace*, *Gaussian* and *Uniform* distributions. Note that for symmetric p.d.f.s all m_n and c_n for odd n are identical to zero and that for the Gaussian distribution all cumulants c_n of order greater than second ($n > 2$) are also zero. Thus second-order statistics $c_2 = m_2 = \sigma^2$,

[6] This treatment can easily be expanded into a multivariate framework (Stein, 1981).

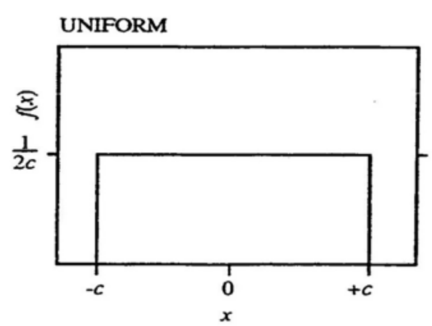
the variance, are sufficient to characterize a Gaussian distribution, since then $c_1 = m_1 = c_3 = m_3 = c_4 = 0$ so that $m_4 = 3m_2^2 = 3\sigma^4$.



n	m_n	c_n
1	0	0
2	2	2
3	0	0
4	24	12



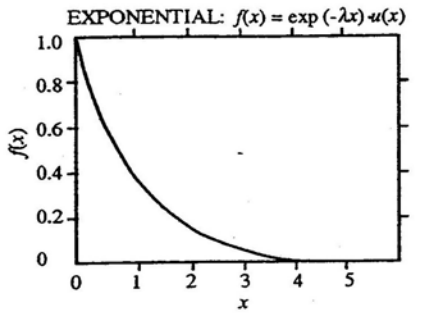
n	m_n	c_n
1	0	0
2	σ^2	σ^2
3	0	0
4	$3\sigma^4$	0



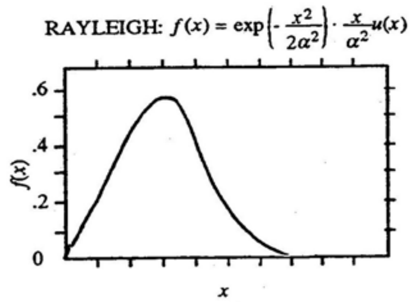
n	m_n	c_n
1	0	0
2	$\frac{c^2}{3}$	$\frac{c^2}{3}$
3	0	0
4	$\frac{c^4}{5}$	$\frac{-2c^4}{15}$

The nth-order moments and cumulants for $n = 1, 2, 3, 4$of the Laplace, Gaussian, and Uniform probability density functions (pdfs)

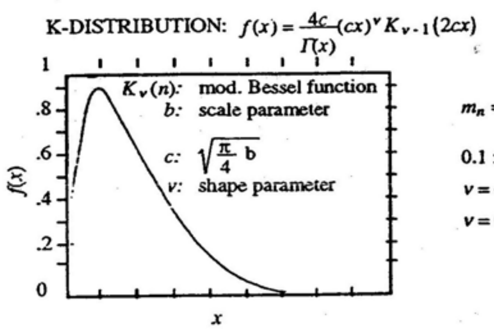
In contrast, Fig. 5 illustrates three non - symmetric parametric distributions: the *Exponential*, *Rayleigh*, and so-called $K-$distributions. It is clear that these distributions require all four order of moments or cumulants to completely describe these distributions.



n	m_n	c_n
1	$\frac{1}{\lambda}$	$\frac{1}{\lambda}$
2	$\frac{2}{\lambda^2}$	$\frac{1}{\lambda^2}$
3	$\frac{6}{\lambda^3}$	$\frac{2}{\lambda^3}$
4	$\frac{24}{\lambda^4}$	$\frac{6}{\lambda^4}$



n	m_n	c_n
1	$\alpha\sqrt{\frac{\pi}{2}}$	$\alpha\sqrt{\frac{\pi}{2}}$
2	$2\alpha^2$	$\left(2-\frac{\pi}{2}\right)\alpha^2$
3	$3\alpha^3\sqrt{\frac{\pi}{2}}$	$-3\alpha^3\pi\sqrt{\frac{\pi}{2}}$
4	$8\alpha^4$	$\frac{1}{2}\alpha^4\left[12\pi - 3\pi^2 - 8\right]$



$$m_n = \frac{1}{c^n} \frac{\Gamma\left(v+\frac{n}{2}\right)}{\Gamma(v)} \cdot \Gamma\left(\frac{n}{2}+1\right)$$

$0.1 \le v \le \infty$

$v = \infty \rightarrow$ Rayleigh

$v = 0.1 \rightarrow$ spiky clutter

The nth-order moments and cumulants for $n = 1, 2, 3, 4$of Exponential, Rayleigh, and K-distribution pdfs

However, moments and cumulants of any order $r > 0$ can be computed for any type of empirical and theoretical distribution, parametric or non-parametric. The fundamental research problem of distributions of empirical financial observations, like speculative stock, bond or foreign exchange prices or their increments, is that we do not know *a priori* how many orders of statistics are sufficient to completely describe such distributions. What can be proved is that, in general, *the computation of joint cumulants of order r requires knowledge of all moments up to order r* (Nikias and Petropulu, 1993).

Furthermore, if a set of variables $\{x(1), x(2), ..., x(n)\}$ is jointly Gaussian, then all the information about their distribution is contained in the moments of order $r \leq 2$. Therefore, all moments of order greater than two ($r > 2$) have no new information to provide. This leads to the fact that all joint cumulants of order $r > 2$ are identical to zero for Gaussian vectors. Hence, the cumulants of order greater than two, in some sense, measure the non - Gaussianness (non-normality) of a distribution. Furthermore, it is always possible to compute the empirical moments and cumulants and then check how close such empirical distributions are to various known parametric distributions. But such approximations by theoretical distributions can never lead to a unique identification of an empirical distribution. There will always be a subjective degree of confidence in the resulting distributional fit.

What do the first four moments measure?

Definition 11 *The **location** of a distribution is measured by the first order statistic, the mean, or average,*

$$c_1 = m_1 = E\{x\} \tag{26}$$

It is easy to *laterally shift* a distribution so that $c_1 = m_1 = 0$ to achieve considerable simplification, by computing deviations from the mean $\varepsilon = x - m_1 = x - E\{x\}$, since then the first order statistic of such deviations ε, $m_1(\varepsilon) = c_1(\varepsilon) = 0$. From now on, we assume that all data are computed as deviations from their means, so that $c_1 = m_1 = 0$.

Definition 12 *The **scale**, dispersion, or variance of a distribution is measured by the second order statistic,*

$$c_2 = m_2 \text{ (since } m_1(\varepsilon) = 0\text{)} \tag{27}$$

Distributions can always be scaled or reduced to normalize distributions, by dividing by the scale. Usually this is done in combination with the lateral shift. Thus, the standardized variable $z = \frac{\varepsilon}{m_2}$, so that $c_1(z) = m_1(z) = 0$ and $m_2(z) = c_2(z) = 1$.

Definition 13 *The* ***skewness*** *of a distribution is measured by its third - order statistic*

$$m_3 = c_3 \text{ (since } m_1(\varepsilon) = 0\text{)} \tag{28}$$

For example, notice in Fig. 4 that the third order statistics of *symmetric* distributions equal zero, $m_3(\varepsilon) = c_3(\varepsilon) = 0$, while those of *asymmetric* distributions are unequal zero, $m_3(\varepsilon) = c_3(\varepsilon) \neq 0$ (since $m_1(\varepsilon) = 0$). In fact, when $m_3(\varepsilon) = c_3(\varepsilon) < 0$ the distribution is *negatively skewed* and when $m_3(\varepsilon) = c_3(\varepsilon) > 0$, the distribution is *positively skewed.*

Definition 14 *The* ***kurtosis,*** *or degree of peakedness, of a distribution is measured by its fourth order statistic*

$$c_4 = m_4 - 3m_2^2 \tag{29}$$

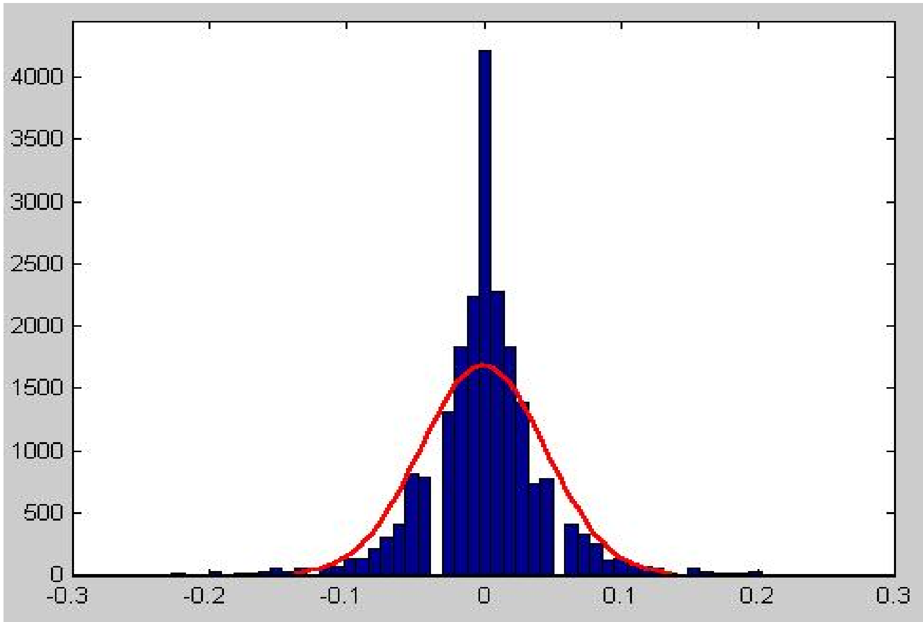
Usually the comparison is made with the Gaussian distribution, which has $m_2 = \sigma^2$, so that we conventionally measure the normalized kurtosis.

Definition 15 ***Normalized kurtosis*** *is measured by*

$$\begin{aligned} \kappa_{normalized} &= \frac{m_4}{m_2^2} \\ &= \frac{c_4}{m_2^2} + 3 \\ &= \frac{c_4}{\sigma^4} + 3 \end{aligned} \tag{30}$$

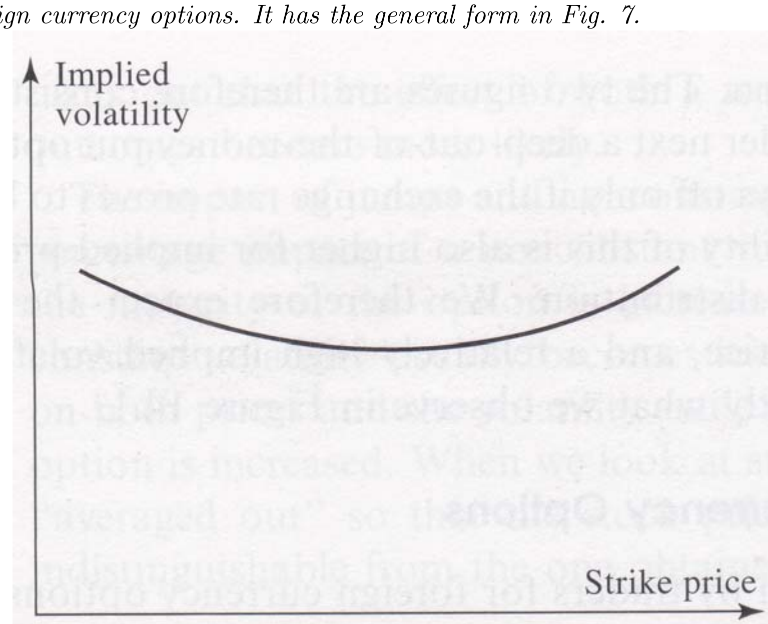
When the kurtosis is the same as that of a Gaussian distribution, $c_4 = 0$, and thus normalized kurtosis equals $\frac{m_4}{m_2^2} = 3$, we speak of **meso - kurtosis**. When $c_4 > 0$, *i.e.*, normalized kurtosis $\frac{m_4}{m_2^2} > 3$, the distribution exhibits large kurtosis, or **lepto - kurtosis**: the frequency distribution is more heavily concentrated around the mean than the Gaussian distribution. When $c_4 < 0$ and normalized kurtosis $\frac{m_4}{m_2^2} < 3$, the distribution exhibits low kurtosis, or **platy - kurtosis**: the frequency distribution is less heavily concentrated about the mean than the Gaussian distribution.

Example 16 *Fig. 6 shows the empirical histogram of the minute - by - minute logarithmic increments of the Japanese Yen for the month of June 1997 compared with the theoretical Gaussian distribution (red line), with the same variance. The distribution of these logarithmic increments is clearly leptokurtic: notice the extreme "peakedness" of the histogram, indicating the higher than normal occurrence of small movements, and the "fat " tails, indicating the higher than normal occurrence of "outliers". There is also a clearly noticeable dearth of intermediate movements.*



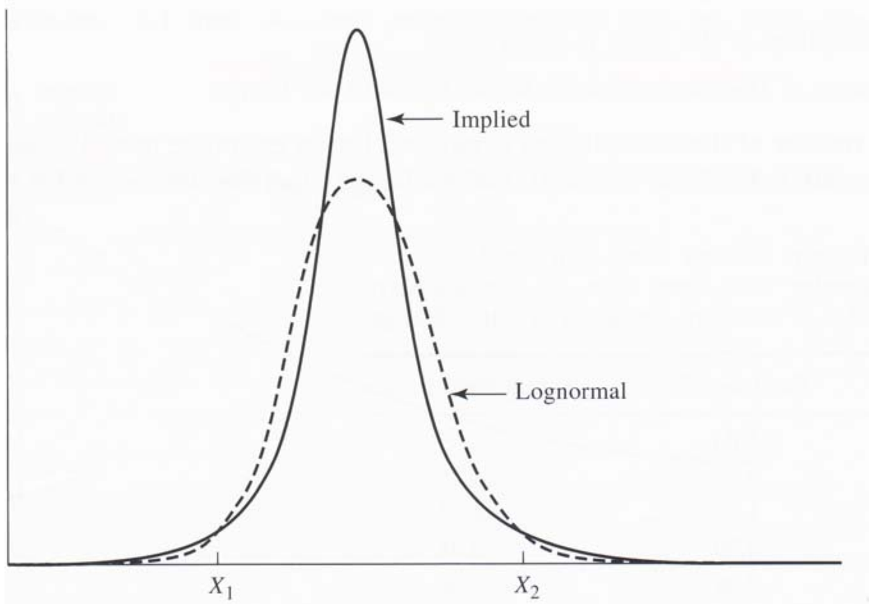
JPY June 1997

Example 17 *A similar phenomenon as in the preceding example has been discussed in the financial literature regarding foreign currency options (cf. Hull, 2001, pp. 286 - 288). The* ***volatility smile****, which relates option volatility σ to the strike price X, is used by traders for empirically pricing of foreign currency options. It has the general form in Fig. 7.*



Volatility smile for foreign currrency options

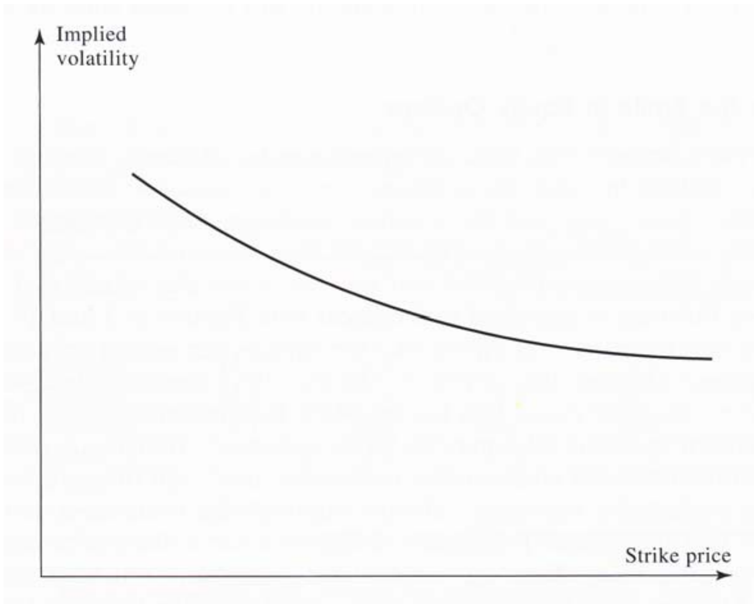
The volatility smile is relatively low for at-the-money options. It corresponds to the (Black - Scholes) implied distribution in Fig. 8, which has higher kurtosis than the corresponding lognormal distribution with the same mean and standard deviation. Figs. 8 and 9 are consistent with each other. Consider first a deep-out-of-the-money call option with a high strike price of X_2. The *relative occurrence of this is higher for the implied distribution than for the lognormal distribution. Therefore, we expect the implied distribution to give a relatively high price for the option. A relatively high option price leads to a relatively high implied volatility.*



Implied distribution and (log-) normal distribution for foreign curency options

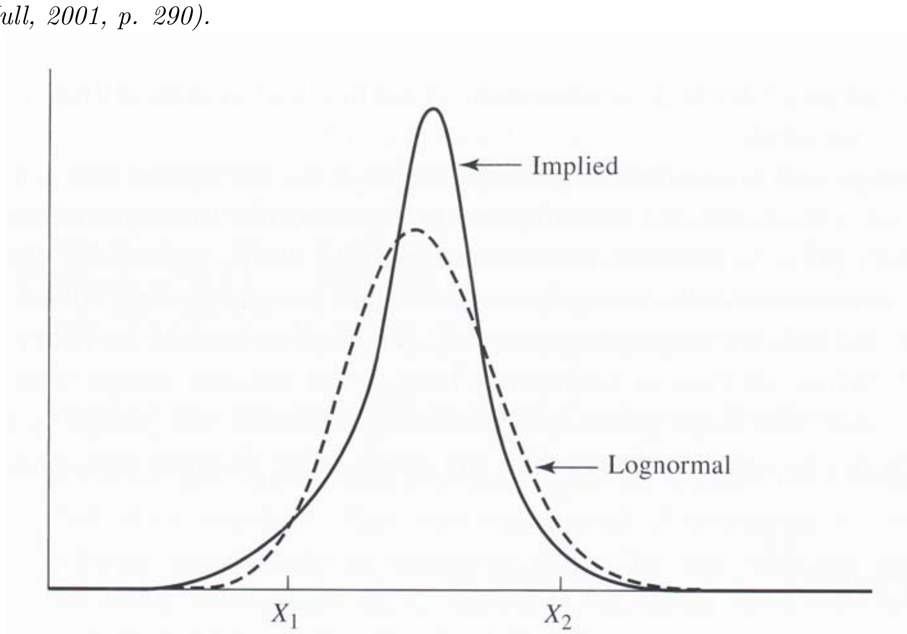
Which is exactly what is empirically observed in Fig. 7. A similar reasoning applies for a deep - out - of - the - money put option with a low strike price of X_1. Thus, the lognormal distribution of foreign currency prices understates the relative occurrence of extreme movements in exchange rates. Based on the measurement of daily movements in 12 different exchange rates over a 10−year period, Hull and White (1998) found that daily changes exceeded three standard deviations on 1.34% *of days, while he lognormal model predicts that this should happen on only* 0.27% *of days. Daily changes exceed four, five and six deviations on* 0.29%, 0.08%, *and* 0.03% *of days, respectively. The lognormal model predicts that we should hardly ever observe this happening. The two reasons for this empirical phenomenon mentioned by Hull and White (1998) are nonconstant volatilities and the impact of jumps (discontinuities) in the exchange rates, often in response to the actions of central banks.*

Example 18 *The traders who empirically price equity options use a **volatility skew**, as in Fig. 9. The volatility σ decreases when the strike price increases (Macbeth and Merville, 1979; Lauterbach and Schultz, 1990; Rubinstein, 1994; Jackwerth and Rubinstein, 1996). The volatility used to price a low strike price option is significantly higher than that used to price a high - strike price option.*



Volatility skew for equities

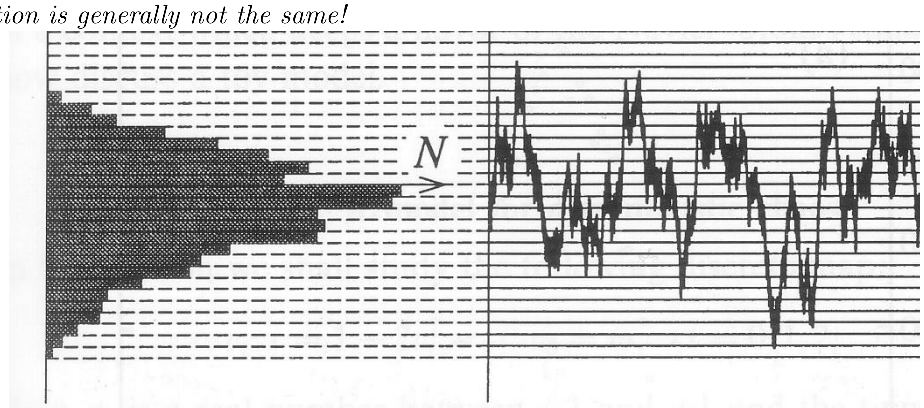
This volatility skew corresponds to the skewed and leptokurtic implied distribution in Fig. 10. It has a fatter left tail and a thinner right tail than the lognormal distribution. For example a deep - out - of - the - money call with a strike price of X_2 has a lower price when the implied distribution is used than when the lognormal distribution is used. A relatively low price leads to a relatively low implied distribution. One possible explanation for this volatility skew in equity options is financial leverage. As a company's equity declines in market value, the company's financial leverage increases. Its equity becomes more risky and its volatility σ increases and vice versa (Hull, 2001, p. 290).



Implied distribution and (log-) normal distribution

Thus we can expect the volatility of equity to be a decreasing function of price, consistent with Figs. 9 and 10, and implying that it is time - dependent. Interestingly, prior to the October 19, *1987 stock market crash implied volatilities were much less dependent on strike prices. Rubinstein (1994) suggests that one reason for the pattern in Fig. 10 may be "crashophobia." Traders are concerned about the possibility of another crash similar to the one of October 19, 1987, and they price options accordingly. It appears that the implied distribution for a stock price has fatter left tails than the distribution calculated from empirical data on stock market returns. Also the volatility skew became more pronounced after the October 1997 and August 1998 declines.*

Example 19 *Fig. 11 shows how an empirical distribution is constructed using a number of bins into which the values of particular ranges are collected. Notice how binning gathers the distributional or frequency information of the time series, but loses the time - dependence information of the series. Until recently, statisticians have paid more attention to the distributional information of time series, while signal processing engineers have paid more attention to their dependence on time. Of course, there is information in both the frequency and the time dimension, and that information is generally not the same!*



Construction of the histogram of a time series by binning

Similarly, Ormoneit and Neumeier (2000, p. 49) show the raw and transformed daily returns of the DAX index in Frankfurt for the period November 1987 - August 1998. The histograms on their right show the relative frequencies of the returns in the same scale.

Example 20 *Table 3 provides the first four moments of the return distributions at different time intervals for the German Deutschemark and the Japanese Yen against the US dollar. The period of observation is January 1, 1987 to June 30, 1996. The data for the USD/DEM and USD/JPY in this Table 3 are selected from Table 2 in Müller, Dacorogna and Pictet (1998), p. 73, which contains similar data for three additional currencies: GBP/USD, USD/CHF, and USD/FRF. The distribution of FX returns is computed from the bid (= intention to buy) and offer (= intention to sell) price quotations of the market maker through the logarithmic middle price*

$$x(t) = \frac{\ln X_{bid,t} + \ln X_{offer,t}}{2} \tag{31}$$

and the return $r(t)$ is measured over a fixed time interval τ as

$$r(t) = x(t) - x(t - \tau) \tag{32}$$

where $x(t)$ is the sequence of logarithmic middle prices spaced equally in Greenwich Mean Time (GMT). The standard deviations are about twice as large as the means and the absolute values of the skewness are mostly significantly smaller than one. From these facts we conclude that the empirical distributions are almost symmetric. The mean values are slightly negative, since during the period of observation there was an overall increase in the value of the USD versus the DEM and versus the JPY (less USD per DEM and JPY, means relatively higher USD value). But comparing these four empirical moments with the theoretical moments of the theoretical distributions of Figs.

4 and 5, we notice that for the shortest time intervals, the measured kurtosis of these empirical distributions is higher than normal ($\gg 3$). Interestingly, all rates show the same general behavior: a decreasing kurtosis with increasing time intervals. At intervals of about one week, the kurtosis is rather close to the Gaussian value. This argues against scaling in the FX markets for time intervals of one week and larger.

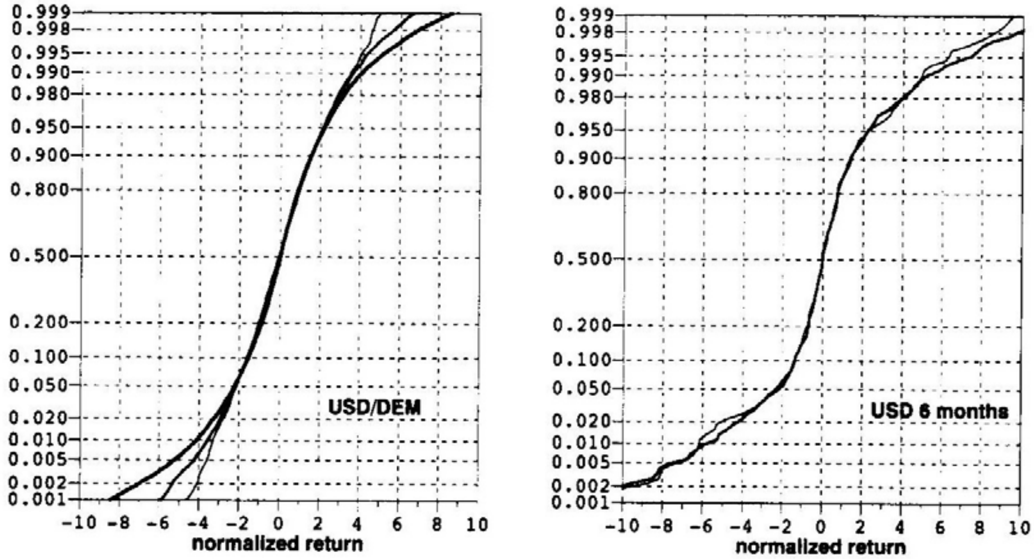
TABLE 3:	***FIRST***	***FOUR***	***MOMENTS***	***OF FX***	***RETURNS***
FX Rate	***Interval τ***	***Mean***	***Variance***	***Skewness***	***Kurtosis***
USD/DEM	*30 minutes*	-1.40×10^{-6}	7.53×10^{-7}	0.60	46.10
	6 hours	-1.68×10^{-5}	8.42×10^{-6}	0.27	11.75
	24 hours	-6.62×10^{-5}	3.48×10^{-5}	0.12	6.04
	1 week	-4.65×10^{-4}	2.41×10^{-4}	0.17	4.24
USD/JPY	*30 minutes*	-2.20×10^{-6}	7.16×10^{-7}	−0.05	24.02
	6 hours	-2.65×10^{-5}	7.98×10^{-6}	−0.17	11.64
	24 hours	-1.06×10^{-4}	3.13×10^{-5}	−0.16	7.06
	1 week	-7.57×10^{-4}	2.22×10^{-4}	−0.23	4.29

Example 21 *Table 4 provides the first four moments of the return distributions at different time intervals for the short term cash interest rates from the interbank money market for the US, Germany, and Japan. The period of observation is January 2, 1979 to June 30, 1996. The data in this Table 4 are selected from Table 3 in Müller, Dacorogna and Pictet (1998), p. 74, which contains similar data for two additional countries, Great Britain and Switzerland. Compare once more these empirical moments with the theoretical moments of the theoretical distributions of Figs. 4 and 5 and notice that these empirical distributions are again not Gaussian, mainly because of their much higher than normal kurtosis ($\gg 3$), but also because of their skewness.*

TABLE 4:	***FIRST***	***FOUR***	***MOMENTS***	***OF CASH***	***RETURNS***
Interest rate	***Interval τ***	***Mean***	***Variance***	***Skewness***	***Kurtosis***
USD 3 months	*24 hours*	-1.27×10^{-5}	2.41×10^{-6}	−0.16	24.72
	1 week	-8.88×10^{-5}	2.20×10^{-5}	−0.53	14.98
USD 6 months	*24 hours*	-1.04×10^{-5}	1.98×10^{-6}	−0.20	20.49
	1 week	-7.33×10^{-5}	1.71×10^{-5}	−0.82	14.48
DEM 3 months	*24 hours*	-1.72×10^{-7}	7.93×10^{-7}	0.39	28.68
	1 week	-7.35×10^{-7}	5.62×10^{-6}	0.22	18.80
DEM 6 months	*24 hours*	-4.76×10^{-7}	7.80×10^{-7}	0.22	33.52
	1 week	-3.99×10^{-6}	5.35×10^{-6}	0.10	11.90
JPY 3 months	*24 hours*	6.06×10^{-7}	1.28×10^{-6}	1.23	43.74
	1 week	4.24×10^{-6}	7.65×10^{-6}	2.80	36.97
JPY 6 months	*24 hours*	-2.47×10^{-6}	9.94×10^{-7}	0.50	46.29
	1 week	-1.73×10^{-5}	6.22×10^{-6}	2.42	28.04

Example 22 *In Fig. 12 three empirical distributions are plotted for the USD/DEM and two for the USD 6 month cash interest rate.*[7] *The cumulative frequency is on the scale of the cumulative Gaussian probability function. Gaussian distributions have the form of a straight line in this representation. Notice that this is approximately the case for the cumulative distribution of weekly returns, whose kurtosis is only slightly higher than normal. In contrast, the distributions of* 30 *- minute and* 24 *- hour returns are distinctly fat-tailed and their kurtosis values are high.*

[7] *This is Figure 2 in Müller, Dacorogna and Pictet (1998), p. 62.*



The empirical cumulative distributions for USD/DEM and USD 6 months cash interest rate, shown for different time intervals τ: 30 minutes, 1 day and 1 week for USD/DEM and 1 day and 1 week for USD 6 months. The fat lines are the shortest time intervals.

But of interest to the research focus of this book, the shape of the FX distribution is not preserved under time aggregation as was the case for the cotton prices in Mandelbrot (1963a). It is clear that the distributions of FX rates are time - varying. In the case of the USD 6 month interest rates, both distributions for the 1− day and 1− week interval look remarkably alike. For these interest rates it is therefore not possible to reject the hypothesis of their distribution being stable under time aggregation. The kurtosis, which is high for the 1−day interval, remains high for the 1−week interval. Also the cumulative distribution for the interest rates is more fat-tailed than for the FX rates, as we can see by comparing Tables 3 and 4. Of course, the empirical distributions of the interest rates are more noisy than for the FX rates, due to both their lower precision of quotation and their lower frequency.

3.2 Stable (Heavy Tailed) Distributions

Stable distributions (or L-stable distributions) are a class of distributions that allow for skewness and heavy tails, like we found in the examples of the cash interest rates. This class of distributions was characterized by Paul Lévy (1954) in his study of sums of independent, identically distributed (i.i.d.) variables. The general stable distribution is described in the Zolotarev parametrization by four parameters, similar to the moments:

(1) an index of stability, or characteristic exponent, $\alpha_Z \in (0, 2]$, which describes the degree of kurtosis;

(2) a skewness parameter $\beta \in [-1, 1]$;

(3) a scale parameter $\gamma > 0$, similar (but not equivalent) to the second distribution moment; and

(4) a location parameter $\delta \in \mathbb{R}$ (a real number). Only a few stable distributions have closed formulas for densities and distribution functions, such as the Gaussian, Cauchy and Lévy distributions.

There are three reasons to use these stable frequency distributions:

(1) There are solid theoretical reasons to expect that real world phenomena exhibit non - Gaussian stable distributions, like in the case of Brownian or Fractional Brownian Motion.

(2) The Generalized Central Limit Theorem states that the only possible non - trivial limit of normalized sums of independent, identically distributed (i.i.d.) variables is stable (and that may be even true for non - i.i.d. variables too. (Cf. Kalman, 1994, 1995).

(3) Empirically, many large data sets exhibit skewness and heavy tails and are poorly described by the Gaussian distribution.

Examples of such stable distributions in finance and economics are given in Mandelbrot (1963, 1966), Fama (1963, 1965), McCulloch (1996), Bassi, Embrechts and Kafetzaki (1998).[8]

4 RANDOM PROCESSES AND TIME SERIES

Thus far we have discussed frequency distributions per se, without taking account of the time - dependence of dynamic financial random processes and of financial time series. In this section we

[8] In 1999 a major AMS - IMS Conference took place, devoted to *Applications of Heavy Tailed Distributions in Economics, Engineering and Statistics*, June 3 - 5, in Washington, DC, USA. The papers of that conference are available on the CD - ROM: *Heavy Tails'99*.

introduce this important time dimension, since most dynamic phenomena exhibit time - dependence characteristics. Of course, we need both a frequency and a time - dependence analysis of a time series, preferably simultaneously, before we can properly model its frequency and dependence characteristics. It is astonishing to observe, how many statisticians, signal processing engineers, and other empirical researchers analyze each of these two characteristics in isolation, or independently from each other, and how often they ignore the other one of these two characteristics.

First, we'll give a formal definitions of a random process and of a time series. Then we discuss some of the peculiar empirical characteristics of financial time series.

Let (Ω, G, P) be a probability space and let T be the ordered set of real numbers corresponding to the times at which the *sequential observations* are carried out

Definition 23 *A monotonically decreasing family of $\sigma-$algebras $\{G_t : t \in T\}$ on a given probability space (Ω, G, P) such that $G_0 \subset G_1 \subset ...... \subset G_{t-1} \subset G_t \subset G$, where G_0 is the trivial algebra $G_0 = \{\emptyset, \Omega\}$, is called a* ***current of $\sigma-$algebras****.*

Definition 24 *The sequence of random variables (r.v.'s) $\{X(t), G_t : t \in T\}$ denotes an object of a current of $\sigma-$algebras $\{G_t : t \in T\}$ on the measurable space (Ω, G), and the sequence of r.v.'s $\{X(t) : t \in T\}$, where the r.v.'s $X(t)$ are G_t-measurable for all $t \in T$, is called a* ***random****, or* ***stochastic process****.*

Some financial series are predictable and some are unpredictable. However, the predictability of a series of observations is not determined by its being random or deterministic. As we already discussed, there exist unpredictable, but deterministic time series, called chaotic series.

Definition 25 *The sequence of r.v.'s $\{X(t) : t \in T\}$ is $\{G_t\}-$**predictable** if $X(t) \in G_{t-1} \subset G_t$ for all $t \in T$.*

There is a more specific, restricted and complete definition of a *dynamical system*, which generalizes the concept of time shifts or time horizons, and which can be used for the analysis of chaos.

Definition 26 *A* ***dynamical system*** *is a quadruplet $(\Omega, \mathcal{A}, P, G)$. The set Ω is the universal space. $\mathcal{A}$ is a $\sigma-$ algebra of Ω. P is the probability measure, which maps $\mathcal{A}$ to the real numbers between 0 and 1 and which satisfies Kolmogorov's axioms*

$$P(A) \geq 0 \text{ for all } A \in \mathcal{A}, \ P(\tag{33}$$

4.1 Stationarity and Serial Dependence

We will now introduce the two essential concepts for time series analysis: stationarity and dependence. These two concepts have been more or less ignored by financial economists and financial analysts, who traditionally, but erroneously, have assumed that all empirical financial and economic time series are stationary and that their increments are mutually independent.

4.1.1 Stationary Processes

Definition 27 *A random process* $\{X(t), G_t : t \in T\}$ *is said to be* ***stationary in the strict sense*** *(distribution stationary, or strongly stationary), if*

$$P\{X(1), X(2), ..., X(t)\} = P\{X(1+\tau), X(2+\tau), ..., X(t+\tau)\} \;\; a.c. \tag{34}$$

for all t *and* $t + \tau \in T$.

Notice that under stationarity in the strict sense, the *whole* joint probability distribution does not change over time. The strict stationarity of random process is equivalent to the property of identically distributiveness (i.d.) of r.v. in classical (non - dynamic) statistics. This is not easy to check empirically, as we will see, since we will have to measure all characteristic moments of the distributions. Regrettably, as we already noted, we can't know *a priori* how many moments are required to characterize an empirical distribution. Thus, pragmatically, we compute only the lower-order moments we're interested in.

Definition 28 *A random process* $\{X(t), G_t : t \in T\}$ *is said to be* ***stationary in the wide sense*** *(covariance stationary, or weakly stationary) if*

$$E\left|X^2(t)\right| < \infty \tag{35}$$

and

$$E\{X(t)X(t-\tau)\} = h(|t-\tau|) \tag{36}$$

i.e., the covariance is a function only of the absolute time period or "lag " $|t - \tau|$.

Stationarity in the wide sense is the one usually assumed since it is empirically much easier to check: it restricts the checking to the second - order moments only.

In finance, this meant that only the first two moments would be computed. But more recently, empirical researchers have become more aware of non - Gaussian distributions and begin to compute the third, fourth and even higher moments, moving beyond the pragmatic checking of stationarity in the wide sense in the direction of the elusive goal of checking for stationarity in the strict sense.

Example 29 *Stationarity in the wide sense is required for classical* ***optimal hedging****, since the use of the minimum variance hedge ratio assumes that the future standard deviations of and the correlation between the changes in the spot and futures prices remain unchanged (Stulz, 1984). The minimum variance hedge ratio is:*

$$h^* = \rho \frac{\sigma_S}{\sigma_F} \tag{37}$$

where σ_S = standard deviations of the changes in the spot price; σ_F = standard deviation of the changes in the futures price; σ_{SF} = covariance between the changes in the spot price and the futures prices, respectively; and $\rho = \frac{\sigma_{SF}}{\sigma_S \sigma_F}$ correlation between spot S and futures F prices. The use of the hedge ratio h^* assumes that σ_S, σ_F and ρ remain time-invariant:

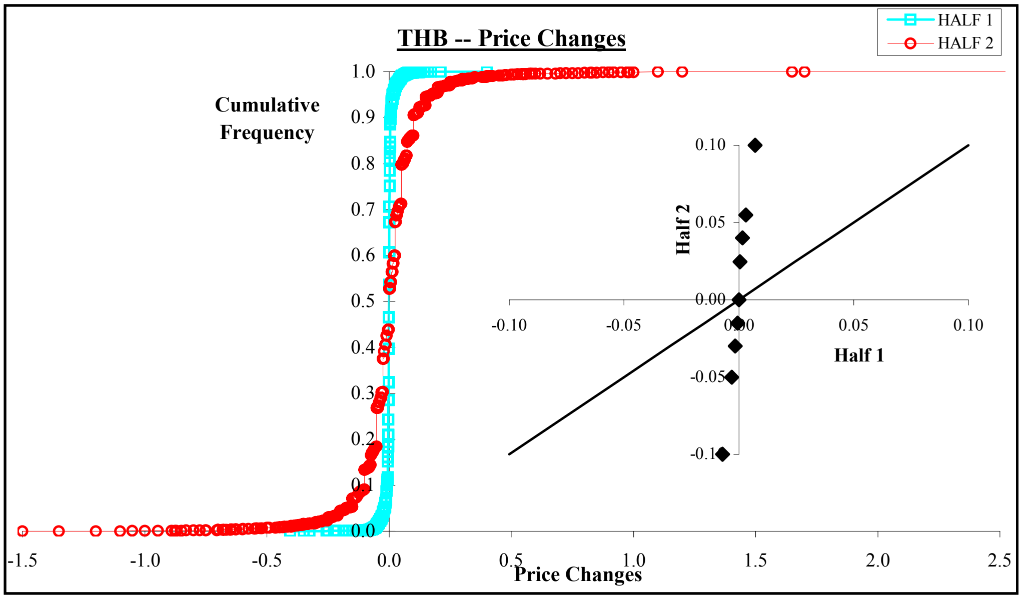
$$\sigma_S, \sigma_F, \rho \;=\; \text{constant for all } t \tag{38}$$

However, the rates of returns of the futures (and spot) markets show little stationarity in the wide sense and, in this context, several questions have been raised about their hedging performance (Ederington, 1979; Franckle, 1980).

It now appears also that, for more complete financial risk measurement, analysis and management, we need to compute at least the first four moments of wide sense stationary processes. Some researchers argue even for several more moments. Next, we must check which of these moments are time - invariant and which are time - varying. It is easy to demonstrate that the first four moments of a simple stock market index series like the S&P500 index are all time - varying and that consequently its distribution is nonstationary in the strict sense.[9]

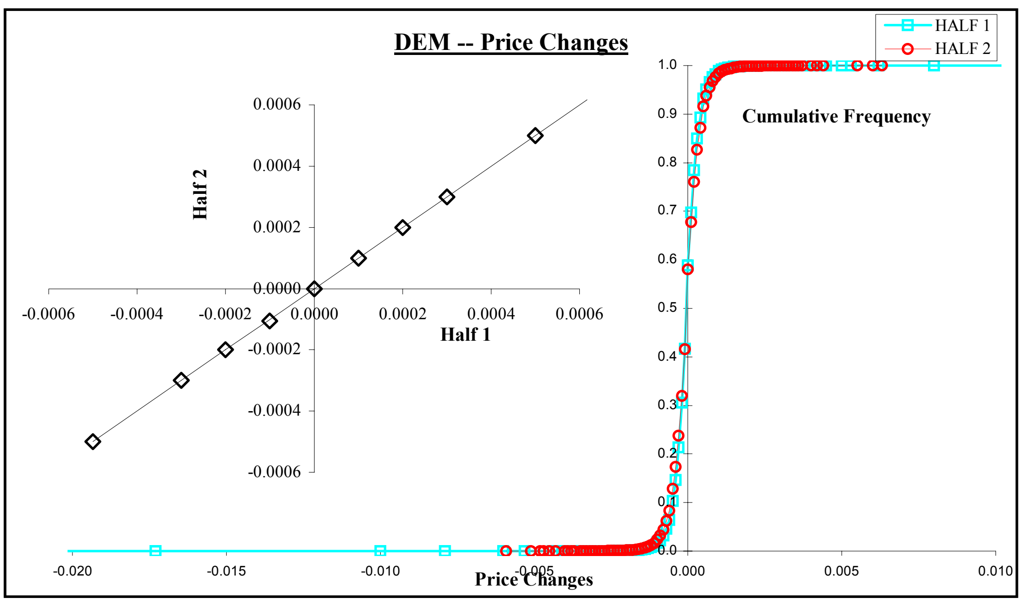
Example 30 *A striking example of a nonstationary foreign exchange rate distribution is given in Fig. 13, which portrays the two semi-annual cumulative distributions of one - minute changes of the Thai baht relative to the US dollar in 1997. The Thai baht abruptly fell on July 2nd, 1997 and the Asian Financial Crisis followed. The distribution of the first half year January - June 1997 in blue is significantly more concentrated than the distribution of the second half year. The small insert chart with differential spectra include data points in both halves corresponding to percentile increments of 10%. Thus the 10, 20, 30%, ..., data values for the second half year are plotted against this in the first half year. Using a 45^0 line to indicate equality between the two half years, the extremely large deviation from the 45^0 line was tested against both halves, and indicated a significant difference (Los, 1999, p. 275; also Abu-Mostafa, LeBaron, Lo and Weigend, 2000, p. 237).*

[9] We do this exercise in other papers together with similar exercises on various financial time series from financial markets in Europe, Latin America and Asia.



Semi-annual Cumultative Distributions of Thai baht - FX Increments, Jan - Dec 1997

In comparison, the shape of the distribution of the Deutschemark, an anchor currency, measured over the same semi - annual periods, remained remarkably stationary in 1997, with the exception of a few outliers, as shown in Fig. 14. (Los, 1999, p. 276; also Abu-Mostafa, LeBaron, Lo and Weigend, 2000, p. 238)



Semi-annual Cumultative Distributions of Deutschemark - FX Increments, Jan - Dec 1997

4.1.2 Moments and cumulants of wide sense stationary processes

The definitions of the cumulants and moments of wide sense stationary processes are straightforward, although not always simple, *e.g.*, the fourth order cumulant is a complex expression (Nikias and Petropulu, 1993, p. 16).

Definition 31 *The first - order cumulant, or **mean value**:*

$$c_1 = m_1 = E\{X(t)\} \tag{39}$$

Definition 32 *Second - order cumulant or **covariance sequence**:*

$$\begin{aligned} c_2(\tau_1) &= m_2(\tau_1) - m_1^2 \\ &= m_2(-\tau_1) - m_1^2 \\ &= c_2(-\tau_1) \end{aligned} \tag{40}$$

where $m_2(\tau_1)$ is the autocorrelation function.

Remark 33 *Notice the time-symmetry of covariance sequences. The time interval is τ_1.*

Remark 34 *When $m_1 = 0$, the covariance sequence simplifies to*

$$c_2(\tau_1) = m_2(\tau_1) \tag{41}$$

Definition 35 ***Third - order cumulant***

$$c_3(\tau_1, \tau_2) = m_3(\tau_1, \tau_2) - m_1[m_2(\tau_1) + m_2(\tau_2) + m_2(\tau_2 - \tau_1)] + 2m_1^3 \tag{42}$$

Remark 36 *When $m_1 = 0$, this simplifies to*

$$c_3(\tau_1, \tau_2) = m_3(\tau_1, \tau_2) \tag{43}$$

Definition 37 ***Fourth - order cumulant***

$$\begin{aligned} c_4(\tau_1, \tau_2, \tau_3) = m_4(\tau_1, \tau_2, \tau_3) - m_2(\tau_1).m_2(\tau_3 - \tau_2) - m_2(\tau_2).m_2(\tau_3 - \tau_1) \\ -m_2(\tau_3).m_2(\tau_2 - \tau_1) - m_1[m_3(\tau_2 - \tau_1, \tau_3 - \tau_1) \\ +m_3(\tau_2, \tau_3) + m_3(\tau_2, \tau_4) + m_3(\tau_1, \tau_2)] \\ +(m_1)^2[m_2(\tau_1) + m_2(\tau_2) + m_2(\tau_3) + m_2(\tau_3 - \tau_1) + m_2(\tau_3 - \tau_2) \\ +m_2(\tau_2 - \tau_1)] - 6(m_1)^4 \end{aligned} \tag{44}$$

Remark 38 *When $m_1 = 0$, this simplifies to*

$$\begin{aligned} c_4(\tau_1, \tau_2, \tau_3) = m_4(\tau_1, \tau_2, \tau_3) - m_2(\tau_1).m_2(\tau_3 - \tau_2) \\ -m_2(\tau_2).m_2(\tau_3 - \tau_1) - m_2(\tau_3).m_2(\tau_2 - \tau_1) \end{aligned} \tag{45}$$

If the random process is zero mean ($m_1 = 0$) [and we can always make it so, if and when the mean is a constant, by analyzing the deviations from the mean], it follows that the second - and third - order cumulants are again identical to the second - and third - order moments, respectively. However, to generate the fourth - order cumulant, we still need knowledge of the fourth- and second - order moments.

4.1.3 Conditional Probability

The study of random processes would not have progressed without the concept of conditional probabilities, which led to so-called *Bayesian* interpretations based on Bayes Theorem.

Definition 39 *The **conditional probability** is*

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \tag{46}$$

which is implicit in Bayes Theorem:

Theorem 40 ***(Bayes)***

$$P(A \mid B)P(B) = P(B \mid A)P(A) = P(A \cap B) \tag{47}$$

Definition 41 *The **conditional expectation** of a random variable $\{X(t) : t \in T\}$ given the $\sigma-$algebra G_t is itself the random variable*

$$E\{X(t) \mid G_t\} \tag{48}$$

measurable with respect to the information set G_t and satisfying the equality

$$\int_A X(t)dP_0 = \int_A E\{X(t) \mid G_t\} dP_0 \tag{49}$$

Examples of wide sense stationary random processes with serial dependence are Markov processes, (arithmetic and geometric) random walks, and nonstationary (G)ARCH processes. Here we look at Markov processes, the grandfather of the other serially dependent random processes to which they all can be reduced.

4.1.4 Markov Process

Definition 42 *The random process $\{X(t) : t \in T\}$, defined on the probability space (Ω, G, P), is said to be a **Markov process in the strict sense**, or to possess the Markov distribution property, if and only if*

$$P\{X(t) \mid X(1), ..., X(t-1)\} = P\{X(t) \mid X(t-1)\} \;\; a.c. \tag{50}$$

Definition 43 *A random process $\{X(t) : t \in T\}$ is said to be a **Markov process in the wide sense** if*

$$E\left\{|X(t)|^2\right\} < \infty \tag{51}$$

and

$$E\{X(t) \mid X(1), ..., X(t-1)\} = E\{X(t) \mid X(t-1)\} \;\; a.c. \tag{52}$$

In general, for a time series, of course

$$E\{X(t) \mid X(t-1)\} \neq X(t-1) \tag{53}$$

Thus a Markov process in the wide sense need not be a martingale.[10] A Markov process in the strict sense involves a stronger restriction than a martingale, since the Markov property involves whole distributions, rather than just the expectations.

4.2 Ergodicity

A crucial, but often misunderstood, or ignored, concept for the classical ensemble approach to time series - which, ideally, considers one historical realization of a time series as only one element of the complete set, or *ensemble,* of possible realizations - is Birkhoff's ergodicity (*cf.* Halmos, 1956 for more details).

Theorem 44 ***(Birkhoff's ergodicity)*** *Let* $\{X(t)\}$ *be a random process, i.e., a measurable, or more specifically, an integrable function defined in the interval* $[0,1]$*. Then the expected value* $E\{.\}$ *(also called the ensemble average) can be replaced by a (limiting) time average, since*

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T}\{X(t).X(t+\tau_1)...X(t+\tau_{r-1})\} = \int_0^1 \{X(t).X(t+\tau_1)...X(t+\tau_{r-1})\} \tag{54}$$

$$= E\{X(t).X(t+\tau_1)...X(t+\tau_{r-1})\} \ \ a.c. \tag{55}$$

If a random process is **ergodic** in the most general form, a.c., all its moments can be determined from a single set of observations. Thus Birkhoff's Theorem allows us to replace time averages over one time interval or orbit by ensemble (or frequency) averages. Clearly, a particular random process might be ergodic for certain higher - order moments but not for others. In practice, when we are given a finite length realization of an ergodic process, we cannot compute the infinite limit, but only the approximating finite estimate

$$\frac{1}{T}\sum_{t=0}^{T}X(t).X(t+\tau_1),...,X(t+\tau_{r-1})$$

Birkhoff's Theorem is the crux of the statistician's conventional frequency oriented approach to time series. It's very controversial and is definitely not accepted by signal engineers on physical grounds. I have not yet seen an empirical scientific test or check for this ideal property of ergodicity,

[10] Fama (1970, 1991) uses the concept of a martingale to define an efficient market. Thus a Markov process can represent efficient and inefficient markets in the Fama sense. But Fama's definition of market efficiency is only true if the empirical pricing processes are wide-sense stationary, which they aren't.

other than the exact computations based on the wavelet scalograms. Based on my observations thus far, I would be very surprised if this equality empirically ever exists and, more important from a scientific perspective, if the assumption of such an equality can ever be checked empirically.

After all, such a scientific check presumes the existence of a complete ensemble of infinitely many possible parallel worlds. But we have only available for scientific analysis one finite length realization, or, at best, a limited number of finite length time realization of a particular dynamic process. The right hand side of the ergodicity equation is based on infinite or complete set of frequencies and the left hand side on an infinite or complete set of time observations. Most empirical invariance properties only hold within limited ranges of finite frequencies and of finite time intervals. Moreover, the Heisenberg Theorem is in direct conflict with the completeness assumptions of Birkhoff's ergodicity.

4.3 Global or Long Term Dependence

The *global dependence*, also known as long term dependence, of random processes and time series, is much harder to define than serial, or short term dependence, such as the wide sense Markov process represents, since in case of global dependence the dependencies are not serial or even overlapping and they may not be linear. Statisticians have the tendency to check for serial correlations = linear dependencies only and not for nonserial, nonlinear dependencies. Global dependence is not of a conditional martingale or Markov nature, since it is not conditional on the whole past, as in the case of martingales, or on the immediately preceding period, as in the case of the Markov property. In fact, global dependence may occur even when serial correlations are close to zero! Global dependence shows correlations only at transient or varying frequencies.

Therefore such global dependencies can better be described in the combined time - frequency domain than in either the time domain or the frequency domain. The frequency domain per se is useful only for characterization of stationary processes. For the characterization of non - stationary, non - linear processes like rates of returns processes in the financial markets one needs to study both their frequency and time characteristics, since the Birkhoff's ergodicity does not exist.[11]

5 CONCLUSIONS

In this essay we argue that the current statistical techniques, which are based on the assumptions of wide - sense stationarity, linearity and ergodicity, are inadequate to analyze the risk characteristics of the pricing processes in the modern financial markets. The analysis of stationary rates of return distributions must (1) be enlarged by including the computation of higher moments and (2) be extended by looking at the non - serial time dependence characteristics of the pricing series. The definition of financial risk should be broader and imply more than the calculation of only a supposedly time-invariant variance or standard deviation for each asset class a, since it is easy to show that such rates of return volatilities are dependent on time t and on the various time horizons τ: $\sigma(a, t, \tau)$. Moreover, higher - order moments of time and horizon dependent skewness and kurtosis must be computed. Finally, financial risk models must be developed beyond the usual serial correlations that take the non - serial, global and non - linear dependencies in rates of return in to account.

6 Bibliography

Alexander, Carol (Ed.) (1999) *Risk Management and Analysis, Volume 1: Measuring and Modelling Financial Risk*, John Wiley & Sons, New York, NY.

Bachelier, L. (1900) "Théorie de la Spéculation," (Doctoral dissertation in Mathematical Sciences, Faculté des Sciences de Paris, defended March 29, 1900), *Annales de l' École Normale Supérieure*, **3**-17, 21 - 86. Translated, with permission of Gauthier-Villars, Paris, France, as Chapter 2 in Cootner, Paul H. (Ed.) (1964) *The Random Character of Stock Market Prices*, The M.I.T. Press, Cambridge, MA, pp. 17 - 78.

[11] In a companion paper "The Changing Concept of Market Efficiency," I spell out in detail the concequences of the changing concept of financial risk for our perception and measurement of financial market efficiency.

Bassi, Franco, Paul Embrechts and Maria Kafetzaki (1998) "Risk Management and Quantile Estimation," in Adler, Robert J., Raisa E. Feldman and Murad S. Taqqu (Eds.) *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, Birkhäuser, Boston, MA, pp. 111 - 130.

Bayes, Thomas (1763) "An Essay Toward Solving a Problem in the Doctrine of Chances," *Philosophical Transactions*, Essay **LII**, 370 - 418. Also in Kendall and Plackett (1977), pp. 134 - 150.

Bernstein, Peter L. (1996) *Against the Gods: The Remarkable Story of Risk*, John Wiley & Sons, Inc., New York, NY.

Bouchaud, Jean-Philippe and Marc Potters (1999) *Theory of Financial Risks: From Statistical Physics to Risk Management*, Cambridge University Press, Chapters 1, 2 and 3 (Translated and expanded from *Théorie des Risques Financiers*, Alea - Saclay - Eyrolles, Paris, 1997).

Chow, Yuan Shih and Henry Teicher (1978) *Probability Theory: Independence, Interchangeability, Martingales*, Springer Verlag, New York, NY.

Cizeau, P., Y. Liu, M. Meyer, C. - K. Peng, and H. Eugene Stanley (1997) "Volatility Distribution in the S&P 500 Stock Index," *Physica*, **A 245**, 441 - 445.

Cowles III, Alfred (1933) "Can Stock Market Forecasters Forecast?", *Econometrica*, **1 -** 3, July, 309 - 324.

Ederington, L. G. (1979) "The Hedging Performance of the New Futures Market," *Journal of Finance*, **34**, March, 157 - 170.

Fama, Eugene F. (1963) "Mandelbrot and the Stable Paretian Hypothesis," *The Journal of Business*, **36** - 4, October, 420 - 429. Reprinted as Chapter 14 in Cootner, Paul H. (Ed.) (1964) *The Random Character of Stock Market Prices*, The M.I.T. Press, Cambridge, MA, pp. 297 - 306.

Fama, Eugene F. (1965a) "The Behavior of Stock-Market Prices," *The Journal of Business*, . **38** - 1, January, 34 - 105.

Fama, Eugene F. (1970) "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, **25**, 383 - 417.

Fama, Eugene F. (1991) “Efficient Capital Markets II, *Journal of Finance*, **46**, 1575 – 1617.

Franckle, C. T. (1980) "The Hedging Performance of the New Futures Market: Comment," *Journal of Finance*, **35**, December, 1273 - 1279.

French, K. and R. Roll (1986) "Stock Return Variances: The Arrival of Information and the Reaction of Traders," *Journal of Financial Economics*, **17**, September, 5 - 26.

Goldreich, Oded (1999) "Pseudo - randomness," *Notices of the American Mathematical Society*, **46 -** 10, November, 1209 - 1216.

Halmos, P. R. (1956) *Lectures on Ergodic Theory*, Chelsea, New York, NY.

Hull, John C. (2001) *Fundamentals of Options and Futures Markets*, Prentice Hall, Upper Saddle River, NJ.

Hull, J. C. and A. White (1998) "Value at Risk When Daily Changes in Market Variables Are Not Normally Distributed," *Journal of Derivatives*, **5** - 3, Spring, 9 - 19.

Ibbotson, Roger G. and Rex A. Sinquefield (1926) *Stocks, Bonds, Bills, and Inflation: 1998 Yearbook (1926 - 1998)*, Dow Jones - Irwin, Chicago, IL.

Jackwerth, J. C. and M. Rubinstein (1996) "Recovering Probability Distributions from Option Prices," *Journal of Finance*, **51**, December, 1611 - 1631.

Kalman, Rudolf E. (1994) "Randomness Reexamined," *Modeling, Identification and Control*, **15-**3, 141 - 151.

Kalman, Rudolf E. (1995) "Randomness and Probability," *Mathematica Japonica*, **41 -** 1, 41 - 58 & "Addendum," **41 -** 2, 463.

Kalman, Rudolf, E. (1996) "Probability in the Real World as a System Attribute," *CWI Quarterly*, **9** - 3, 181 - 204.

Kendall, Maurice G. and R. L. Plackett (Eds.) (1977) *Studies in the History of Statistics and Probability*, Vol. II, Macmillan, New York, NY.

Keynes, John Maynard (1921) *A Treatise on Probability*, Macmillan, London, UK.

Kindleberger, Charles P. (1996) *Manias, Panics, and Crashes: A History of Financial Crises*, John Wiley & Sons, New York, NY.

Knight, Frank H. (1964) *Risk, Uncertainty & Profit*, Century Press, New York, NY (Original: 1921).

Kolmogorov, A. N. (1933) *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer, Berlin (Translated by Nathan Morrison: *Foundations of Probability*, Chelsea, New York, NY, 1950).

Kondor, I. and J. Kertesz (Eds.) (1999) *Econophysics: an Emerging Science*, Kluwer, Dordrecht, The Netherlands.

Lauterbach, B. and P. Schultz (1990) "Pricing Warrants: An Empirical Study of the Black - Scholes Model and Its Alternatives," *Journal of Finance*, **44**, September, 1181 - 1210.

Los, Cornelis A. (1982) "Discrete-Time Martingale Convergence Results and Nonlinear Estimation Using Time-Series Data," Research Paper No: 8222, Federal Reserve Bank of New York, New York, 107 pages.

Los, Cornelis A. (1998) "Optimal Multi - Currency Investment Strategies With Exact Attribution in Three Asian Countries," *Journal of Multinational Financial Management*, **8** - 2 -3, September, 169 - 198

Los, Cornelis A. (1999) "Nonparametric Testing of the High-Frequency Efficiency of the 1997 Asian Foreign Exchange Markets," *Journal of Multinational Financial Management*, **9-**. 3 - 4, October, 265 - 289.

Los, Cornelis A. (2000a) "Nonparametric Efficiency Testing of Asian Stock Markets, Using Weekly Data," in Fomby, Thomas, B. and R. Carter Hill (Editors), *Advances in Econometrics: Applying Kernel and Nonparametric Estimation to Economic Topics*, **14**, JAI Press, Inc., pp. 329 - 363.

Los, Cornelis A. (2000b), "Frequency and Time Dependence of Financial Risk," *The Journal of Performance Measurement*, **5** - 1, Fall, 72 - 73.

Los, Cornelis A. (2001) *Computational Finance: A Scientific Perspective*, World Scientific Publishing, Co., Singapore.

Macbeth, J. D. and L. J. Merville (1979) "An Empirical Investigation of the Black - Scholes Call Option Pricing Model," *Journal of Finance*, **34**, 1173 - 1186.

Mandelbrot, Benoit B. (1963a) "New Methods in Statistical Economics," *Journal of -Political Economy*, **71**, 421 - 440. Reprinted as Chapter E3, with annotations and corrections, in *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk*, Springer Verlag, New York, NY, 1997, 471 - 491.

Mandelbrot, Benoit B. (1963b) The Variation of Certain Speculative Prices, *The Journal of Business*, **36**, 394 - 419 & **45**, 1972, 542 - 543. Reprinted as Chapter E14, with annotations and corrections, in *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk*, Springer Verlag, New York, NY, 1997, 371 - 418. Also in Cootner, 1964, 297 - 337.

Mandelbrot, Benoit B. (1966) "Forecasts of Future Prices, Unbiased Markets and Martingale Models," *The Journal of Business*, **39**, January (Special Supplement), 242 - 255. Reprinted as Chapter E19 "Nonlinear Forecasts, Rational Bubbles, and Martingales," with annotations and corrections, in *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk*, Springer Verlag, New York, NY, 1997, 471 - 491.

Mandelbrot, B. B. and J. R. Wallis (1969) "Some Long-run Properties of Geophysical Records," *Water Resources Research*, **5**, 321 - 340.

Mantegna, Rosario, N. (Ed.) (1999) *Proceedings of the International Workshop on Europhysics and Statistical Finance*, *Physica*, **A 269** (Special Issue).

McCulloch, J. Huston. (1996) "Financial Applications of Stable Distributions," in Maddala, G. S. and C. R. Rao (Eds.), *Statistical Methods in Finance* (*Handbook of Statistics*, **14**, Elsevier, Amsterdam, The Netherlands, pp. 393 - 425.

Merton, Robert C. (1999) *Continuous Time Finance*, revised edition, Basil Blackwell, Oxford, UK.

Meyer, Yves (1987) "Principe d'Incertitude, Bases Hilbertienne et Algèbres d'Opérateurs" ("Uncertainty Principle, Hilbert bases and Operator Algebra"), in *Séminaire Bourbaki*, **145 -**

146, 209 - 223, Astérisque, Paris.

Mittnik, Stefan, Svetlozar T. Rachev and Marc S. Paolella (1998) "Stable Paretian Modeling in Finance," in Adler, Robert J., Raisa E. Feldman and Murad S. Taqqu (Eds.) (1998) *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, Birkhäuser, Boston, MA, pp. 79 - 110.

Müller, Ulrich A., Michel A. Dacorogna and Olivier V. Pictet (1998) "Heavy Tails in High - Frequency Financial Data," in Adler, Robert J., Raisa E. Feldman and Murad S. Taqqu (Eds.) *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, Birkhäuser, Boston, MA, pp. 55 - 77.

Nikias, Chrysostomos L. and Athina P. Petropulu (1993) *Higher - Order Spectra Analysis: A Nonlinear Signal Processing Framework*, PTR Prentice Hall, Englewood Cliffs, NJ.

Papoulis, A. (1984) *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, New York, NY.

Peters, Edgar E. (1999) *Patterns in the Dark: Understanding Risk and Financial Crisis with Complexity Theory*, John Wiley & Sons, New York, NY.

Pincus, Steve and Rudolf E. Kalman (1997) "Not All (Possibly) "Random" Sequences Are Created Equal," *Proceedings of the National Academy of Sciences (USA)*, **94**, April, 3513 - 3518.

Pincus, S. and B. H. Singer (1996) "Randomness and Degrees of Irregularity," *Proceedings of the National Academy of Sciences (USA)*, **93**, 2083 - 2088.

Mantegna, Rosario, N. and H. Eugene Stanley (2000) *An Introduction to Econophysics: Correlations and Complexity in Finance*, Cambridge University Press, Cambridge, UK.

Rubinstein, M. (1994) "Implied Binomial Trees," *Journal of Finance*, **49** - 3, July, 771 - 818.

Savit, Robert (1988) "When Random is Not Random: an Introduction to Chaos in Market Prices," *Journal of Futures Markets*, **8**, 271.

Savit, Robert (1994) "Chaos on the Trading Floor," Chapter 14 in Hall, Nina (Ed.) *Exploring Chaos: A Guide to the New Science of Disorder*, W. W. Norton & Company, New York, NY, 1994, 174 - 183.

Schroeder, Manfred (1999) *Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise*, W. H. Freeman and Co., New York, NY.

Schwert, G. W. (1989) "Why Does Stock Market Volatility Change over Time?", *Journal of Finance*, **44**, 1115 - 1153.

Stein, C. (1980) "Estimation of the Mean of a Multivariate Normal Distribution," *Annals of Statistics*, **10**, 1135 - 1151.

Stulz, R. M. (1984) "Optimal Hedging Policies," *Journal of Financial and Quantitative Analysis*, **19**, 127 - 140.

Tobin, James (1958) "Liquidity Preference as Behavior Towards Risk," *Review of Economic Studies*, **25** - 2, February, 65 - 85.