ARCH and GARCH models:
description of stochastic processes with a time-dependent variance

(Chapter 10 of Mantegna–Stanley book)

István Szanyi

Econophysics Meeting
November 21, 2024

Probability space & stochastic process

- **sample space** (Ω): the set of all possible outcomes
- **event space** (F): a collection of events to be considered (subsets of Ω)
- **probability measure** (P): a function that returns an event's probability

The **probability space** is the mathematical triplet (Ω, F, P) that presents a model for a particular class of real-world situations.

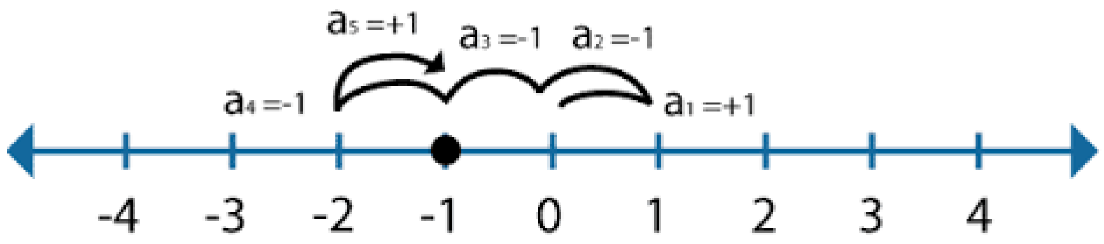
A **stochastic or random process** $x(t)$ is a collection of random variables $\{x_t\}$ with index t defined on a common probability space (Ω, F, P).

Stochastic process

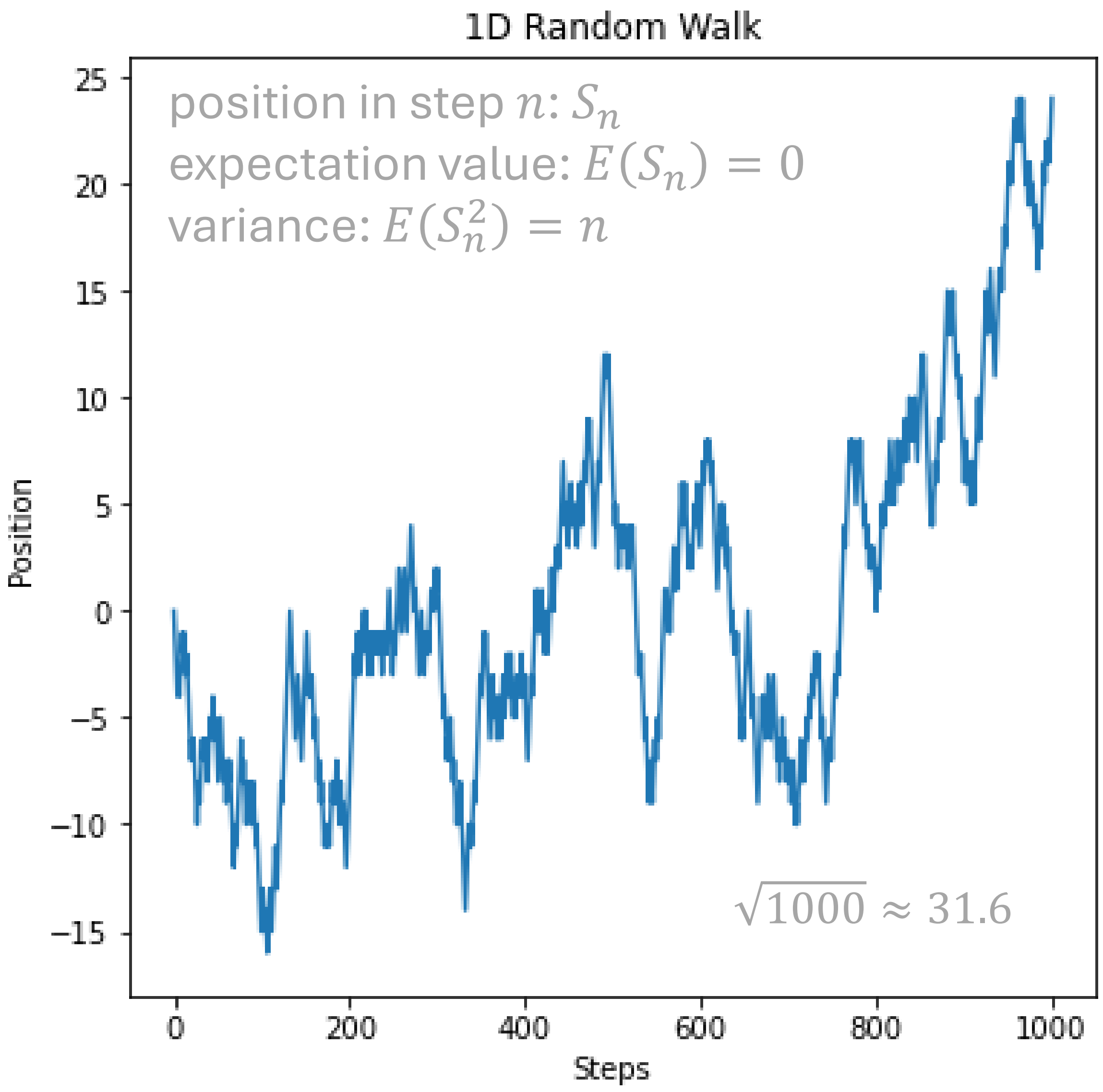
- a collection of random variables
- random variables are indexed by some mathematical set, the **index set** (each random variable of the stochastic process is uniquely associated with an element in that set, e.g., $\mathbb{N}$)
- each random variable in the collection takes values from the same mathematical space known as the **state space** (e.g., $\mathbb{Z}$)
- **increment**: the amount that a stochastic process changes between two index values (e.g., between two points in time).
- **sample function** or **realization**: a single outcome of a stochastic process

Random walk as a stochastic process

X_1 , X_2 , … are independent random varaibles, where each variable is either 1 or −1, with a 50% probability for either value; $S_0 = 0$ and $S_n = \sum_{j=0}^{n} X_j$. The series $\{S_n\}$ is called the **simple random walk** on $\mathbb{Z}$.



- sample space: $\Omega = \{-1,1\}$
- event space: $F = \{-1,1\}$
- probability measure: P; $P(-1) = 1/2$, $P(1) = 1/2$
- index set: $\mathbb{N}$
- state space: $\mathbb{Z}$
- increment: -1 or 1



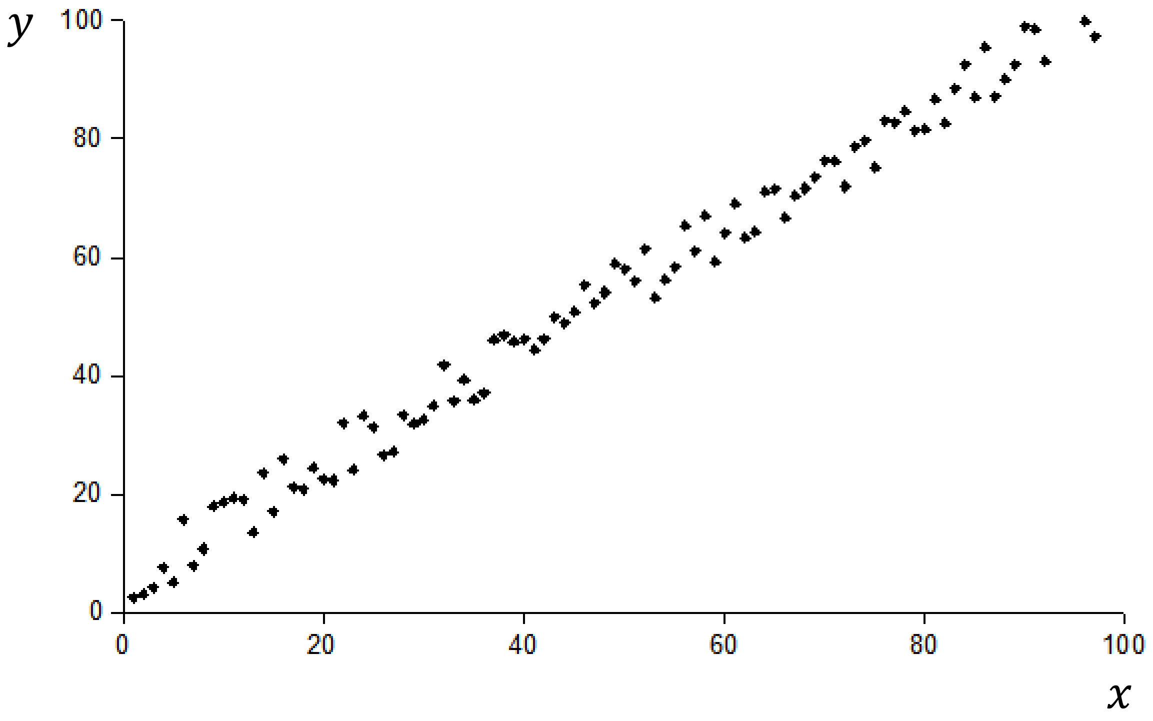
Sample function / realization of a one-dimensional simple random walk (n=1000).

Homoscedasticity & heteroscedasticity

"Skedasticity" comes from the Ancient Greek word "skedánnymi", meaning "to scatter".

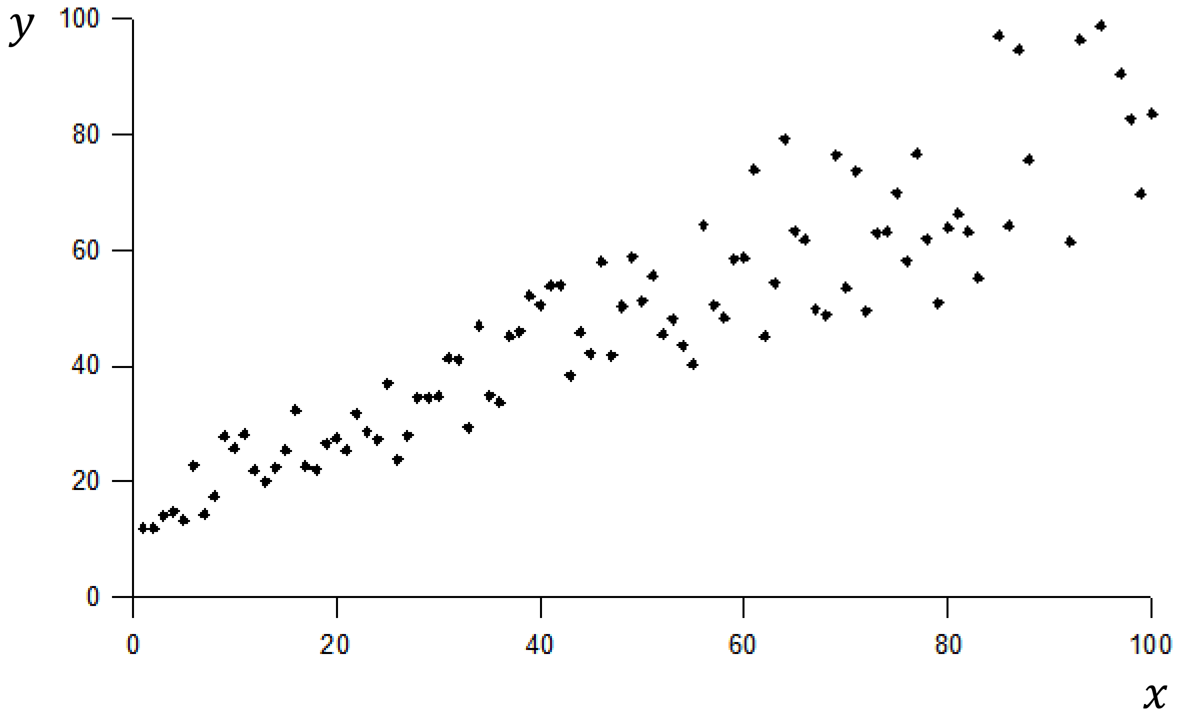
homoscedasticity

at each value of *x*, the *y*-value of the data points has about the *same variance*



heteroscedasticity

the *variance* of the *y*-values of the dots *increases* with increasing values of *x*.



ARCH processes/models

A stochastic process *with nonconstant variances conditional on the past, but constant unconditional variances, i.e., with **autoregressive conditional heteroskedasticity*** is an **ARCH(p) process** of order p.

ARCH models are discrete-time stochastic models for which the variance at time t, *i.e.*, σ_t^2 depends, conditionally, on some past values of the square value of the random signal itself.

$$\sigma_t^2 = \alpha_0 + \alpha_1 x_{t-1}^2 + \cdots + \alpha_p x_{t-p}^2. \tag{10.1}$$

Here $\alpha_o, \alpha_1, \ldots, \alpha_p$ are positive variables and x_t is a random variable with zero mean and variance σ_t^2, characterized by a conditional pdf $f_t(x)$. Usually $f_t(x)$ is taken to be a Gaussian pdf, but other choices are possible.

Conditional distributions, memory

$$\boxed{\Pr\left(X_{t_k} | X_{t_{k-1}}, X_{t_{k-2}}, \ldots, X_{t_1}\right)}$$

for some set of times $t_k > t_{k-1} > \ldots > t_1$. In general, this conditional distribution will depend upon values of $X_{t_{k-1}}, X_{t_{k-2}}, \ldots, X_{t_1}$. However, we shall focus particularly in this module on processes that satisfy the **Markov** property, which says that

$$\boxed{\Pr\left(X_{t_k} | X_{t_{k-1}}, X_{t_{k-2}}, \ldots, X_{t_1}\right) = \Pr\left(X_{t_k} | X_{t_{k-1}}\right).}$$

The Markov property is named after the Russian probabilist Andrei Andreyevich Markov (1856-1922). An informal mnemonic for remembering the Markov property is this. 'Given the present (X_{k-1}), the future (X_k) is independent of the past ($X_{k-2}, X_{k-3}, \ldots, X_1$).' The Markov property is sometimes referred to as the 'lack of memory' property.

ARCH processes/models

$$\sigma_t^2 = \alpha_0 + \alpha_1 x_{t-1}^2 + \cdots + \alpha_p x_{t-p}^2. \tag{10.1}$$

By varying the number p of terms in Eq. (10.1), one can control the amount and the nature of the memory of the variance σ_t^2. Moreover, the stochastic nature of the ARCH(p) process is also changed by changing the form of the conditional pdf $f_t(x)$. An ARCH(p) process is completely determined only when p and the shape of $f_t(x)$ are defined.

ARCH models are simple models able to describe a stochastic process which is locally nonstationary but asymptotically stationary: **the parameters controlling the conditional probability density function $f_t(x)$ at time t are fluctuating; the stochastic process has a well-defined asymptotic PDF $P(x)$**

Numerical simulations of ARCH(1) processes

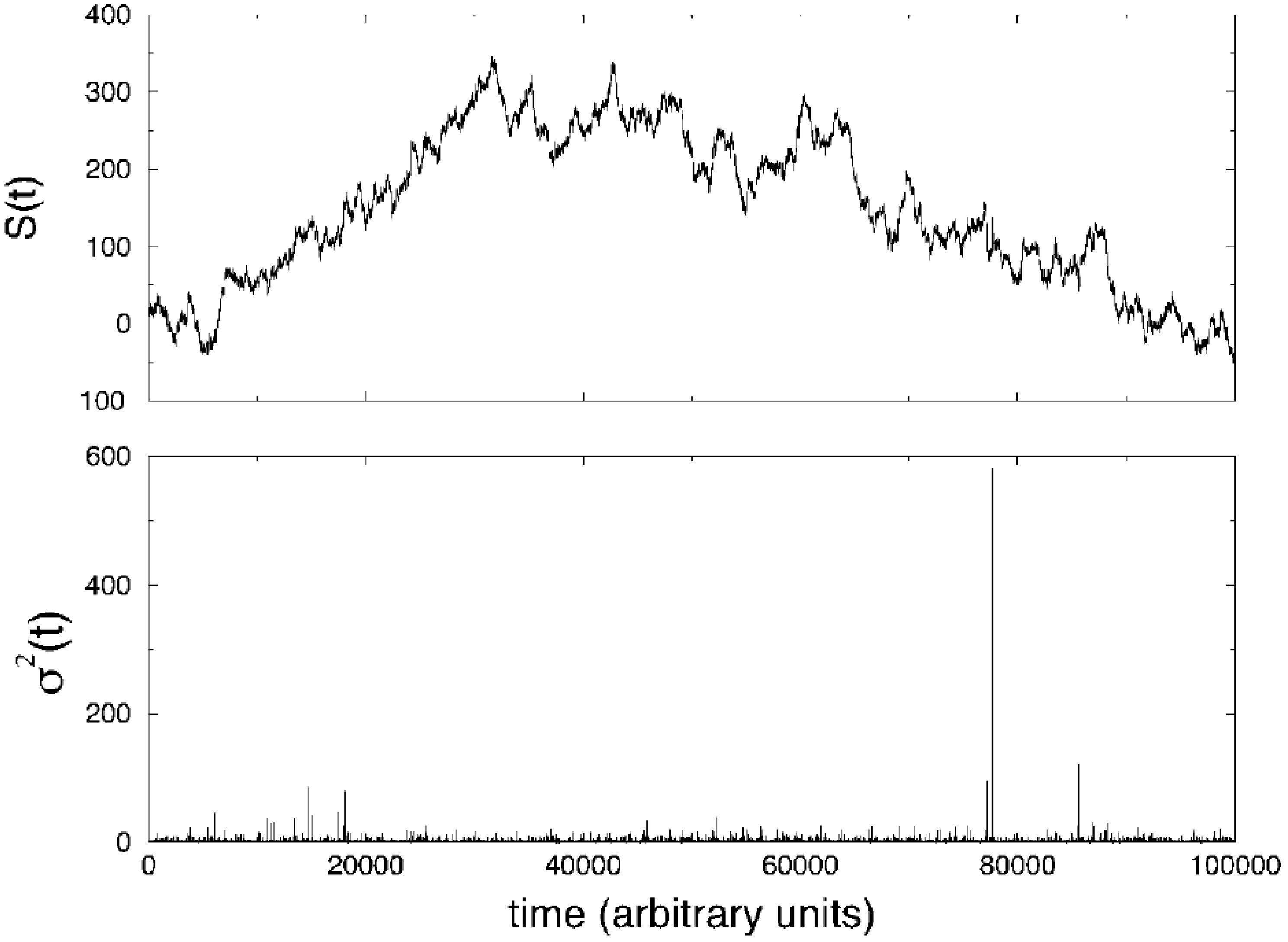
simplest ARCH process, ARCH(1) process with Gaussian conditional PDF:

$$\sigma_t^2 = \alpha_0 + \alpha_1 x_{t-1}^2$$

$$S(t) = \sum_{i=1}^{i=t} x_i.$$

$$\alpha_0 = 0.45 \text{ and } \alpha_1 = 0.55$$

Although the conditional pdf is chosen to be Gaussian, the asymptotic pdf presents a given degree of leptokurtosis (fatter tails) because the variance σ_t of the conditional pdf is itself a fluctuating random process.



Asymptotic ‘unconditional’ variance, kurtosis

$$\boxed{\sigma_t^2 = \alpha_0 + \alpha_1 x_{t-1}^2}$$

An ARCH(1) process with Gaussian conditional pdf is characterized by a finite ‘unconditional’ variance (the variance observed on a long time interval), provided

$$1 - \alpha_1 \neq 0 \qquad 0 \leq \alpha_1 < 1. \tag{10.4}$$

The value of the variance is

$$\sigma^2 = \frac{\alpha_0}{1 - \alpha_1}. \tag{10.5}$$

The kurtosis of the ARCH(1) process is [50]

$$\kappa \equiv \frac{\langle x^4 \rangle}{\langle x^2 \rangle^2} = 3 + \frac{6\alpha_1^2}{1 - 3\alpha_1^2}, \tag{10.6}$$

which is finite if

$$0 \leq \alpha_1 < \frac{1}{\sqrt{3}}. \tag{10.7}$$

Hence, by varying α_0 and α_1, it is possible to obtain stochastic processes with the same unconditional variance but with different values of the kurtosis.

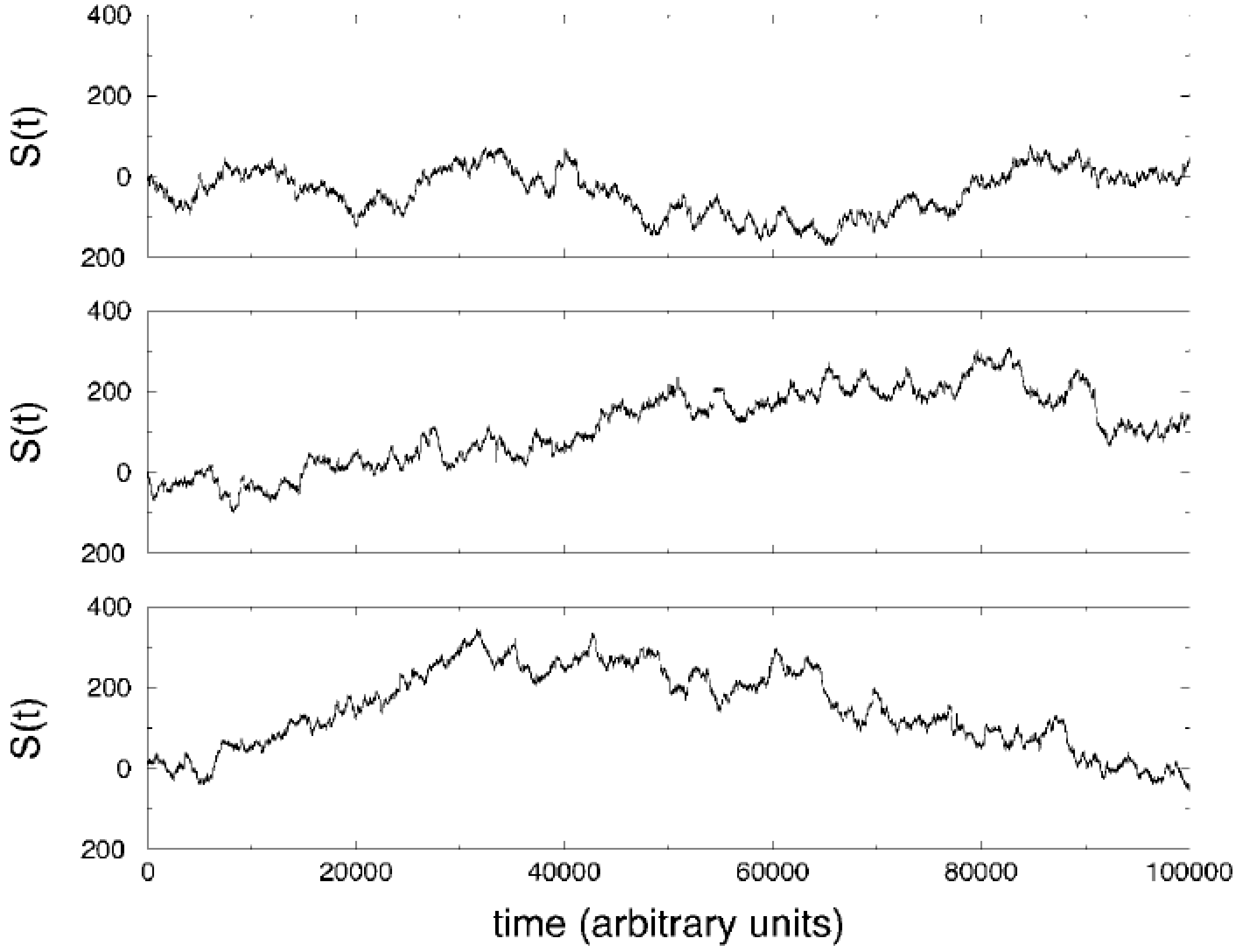


Fig. 10.2. Numerical simulations of ARCH(1) processes with the same unconditional variance ($\sigma^2 = 1$) and different values of the unconditional kurtosis. Top: $\alpha_0 = 1$, $\alpha_1 = 0$ (so $\kappa = 3$ by Eq. (10.6)). Middle: $\alpha_0 = \alpha_1 = 0.5$ (so $\kappa = 9$). Bottom: $\alpha_0 = 0.45$, $\alpha_1 = 0.55$ (so $\kappa = 23$).

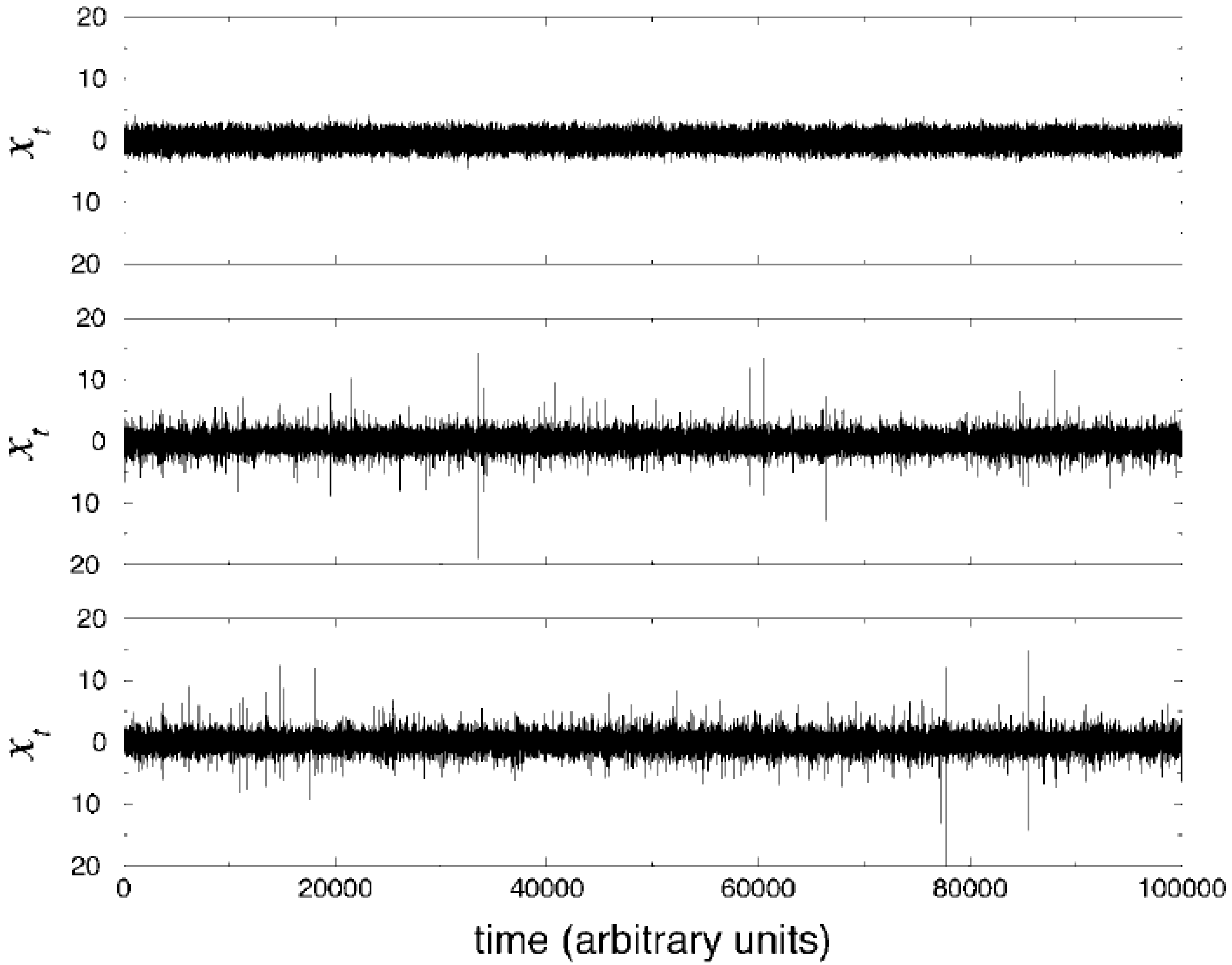


Fig. 10.3. Successive increments of the simulations shown in Fig. 10.2. Events outside three standard deviations are almost absent when $\kappa = 3$ (top), are present when $\kappa = 9$ (middle), and are more intense when $\kappa = 23$ (bottom).

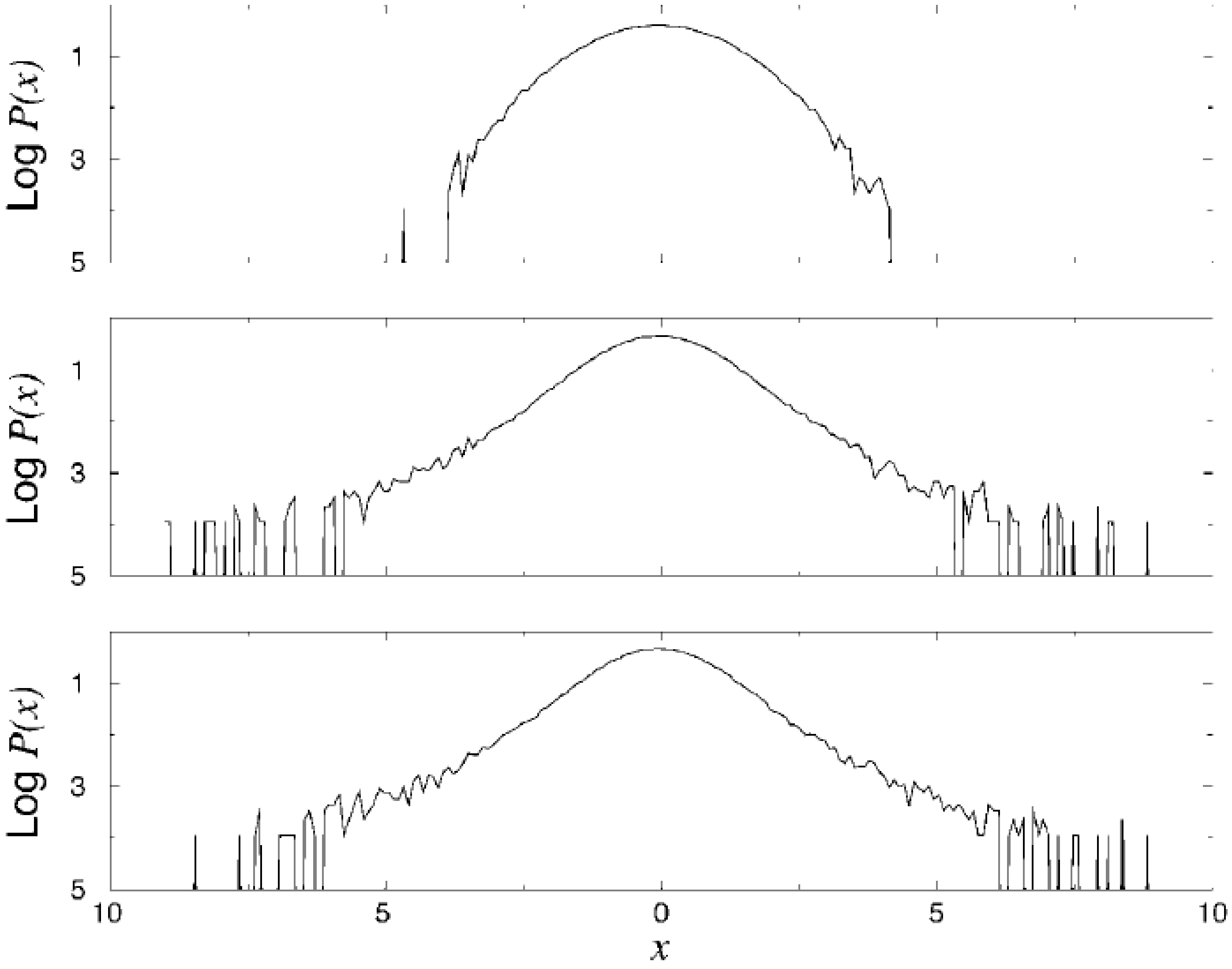


Fig. 10.4. Probability density function of the successive increments shown in Fig. 10.3. The pdf is Gaussian when $\kappa = 3$ (top) and is leptokurtic when $\kappa = 9$ or 23 (middle and bottom).

GARCH processes

In many applications using the linear ARCH(p) model, a large value of p is required. This usually poses some problems in the optimal determination of the $p+1$ parameters $\alpha_0, \alpha_1, \ldots, \alpha_p$, which best describe the time evolution of a given economic time series. The overcoming of this difficulty leads to the introduction of generalized ARCH processes, called GARCH(p, q) processes, introduced by Bollerslev in 1986 [20].

This class of stochastic processes is defined by the relation

$$\boxed{\sigma_t^2 = \alpha_0 + \alpha_1 x_{t-1}^2 + \cdots + \alpha_q x_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \cdots + \beta_p \sigma_{t-p}^2,} \tag{10.8}$$

where $\alpha_0, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p$ are control parameters. Here x_t is a random variable with zero mean and variance σ_t^2, and is characterized by a conditional pdf $f_t(x)$, which is arbitrary but is often chosen to be Gaussian.

We consider the simplest GARCH process, namely the GARCH(1,1) process, with Gaussian conditional pdf. It can be shown [9] that

$$\sigma^2 = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}, \tag{10.9}$$

and the kurtosis is given by the relation

$$\kappa = 3 + \frac{6\alpha_1^2}{1 - 3\alpha_1^2 - 2\alpha_1\beta_1 - \beta_1^2}. \tag{10.10}$$

Mean and autocorrelation

The statistical observables characterizing a stochastic process can be written in terms of nth-order statistical properties. The case $n = 1$ is sufficient to define the mean,

$$E\{x(t)\} \equiv \int_{-\infty}^{\infty} x f(x,t)dx, \tag{6.1}$$

where $f(x,t)$ gives the probability density of observing the random value x at time t. The case $n = 2$ is used to define the autocorrelation function

$$\boxed{E\{x(t_1)x(t_2)\} \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f(x_1, x_2; t_1, t_2)dx_1 dx_2,} \tag{6.2}$$

where $f(x_1, x_2; t_1, t_2)$ is the joint probability density that x_1 is observed at time t_1 and x_2 is observed at time t_2. To fully characterize the statistical properties of a stochastic process, knowledge of the function $f(x_1, x_2, \ldots, x_n; t_1, t_2, \ldots, t_n)$ is required for every x_i, t_i and n. Most studies are limited to consideration of the 'two-point' function, $f(x_1, x_2; t_1, t_2)$.

$$\boxed{R(t_1, t_2) = R(\tau) \text{ is a function of } \tau \equiv t_2 - t_1}$$

Autocorrelation, also known as serial correlation, refers to the degree of correlation of the same variables between two successive time intervals. The value of autocorrelation ranges from -1 to 1.

Memory

Now we focus on the kind of time memory that can be observed in stochastic processes. An important question concerns the typical scale (time memory) of the autocorrelation function. For stationary processes, we can answer this important question by considering the integral of $R(\tau)$. The area below $R(\tau)$ can take on three possible values (Fig. 6.2),

$$\int_0^\infty R(\tau)d\tau = \begin{cases} \text{finite} \\ \text{infinite} \\ \text{indeterminate} \end{cases} . \tag{6.9}$$

When $\int_0^\infty R(\tau)d\tau$ is finite, there exists a typical time memory τ_c called the correlation time of the process.

Properties of GARCH(1,1) with Gaussian conditional pdf

$$\boxed{\sigma_t^2 = \alpha_0 + \alpha_1 x_{t-1}^2 + \beta_1 \sigma_{t-1}^2.} \tag{10.11}$$

The random variable x_t can be written in term of σ_t by defining

$$x_t \equiv \eta_t \sigma_t, \tag{10.12}$$

where η_t is an i.i.d. random process with zero mean, and unit variance. Under the assumption of Gaussian conditional pdf, η_t is Gaussian. By using Eq. (10.12), one can rewrite Eq. (10.11) as

$$\sigma_t^2 = \alpha_0 + (\alpha_1 \eta_{t-1}^2 + \beta_1)\sigma_{t-1}^2. \tag{10.13}$$

Equation (10.13) shows that GARCH(1,1) and, more generally, GARCH(p, q) processes are essentially random multiplicative processes. The autocorrelation function of the random variable x_t, $R(\tau) = \langle x_t x_{t+\tau} \rangle$ is proportional to a delta function $\delta(\tau)$.

GARCH(1,1) as a Markovian process

What about the higher-order correlation of the process? Following Bollerslev [20], we will see that in a GARCH(1,1) process, x_t^2 is a Markovian random variable characterized by the time scale $\tau = |\ln(\alpha_1 + \beta_1)|^{-1}$. Hence a GARCH(1,1) process provides an interesting example of a stochastic process x_t that is second-order uncorrelated, but is higher-order correlated.

From Eq. (10.17), we see that the autocovariance of the square of the process x_t is described by the exponential form

$$\mathrm{cov}(x_t^2, x_{t+n}^2) = A\, e^{-n/\tau}, \tag{10.20}$$

where $A \equiv \alpha_0/(1 - B)$ and $\tau \equiv |\ln B|^{-1}$, and $B \equiv \alpha_1 + \beta_1$. In a GARCH(1,1) process the square of the process x_t^2 is a Markovian process characterized by the time scale τ.

Properties of GARCH(1,1) with Gaussian conditional pdf

the variance of returns is characterized by a power-law correlation. Since the correlation of the square of a GARCH(1, 1) process is exponential, a GARCH(1, 1) process cannot be used to describe this empirically observed phenomenon properly. In spite of this limitation, GARCH(1, 1) processes are widely used to model financial time series. The limitation is overcome by using values of B close to one in empirical analysis [1]. Values of B close to one imply a time memory that could be of the order of months. The model's values for the α_1 and β_1 parameters – obtained in the period 1963 to 1986 by analyzing the daily data of stock prices of the Center for Research in Security Prices (CRSP) – give $\alpha_1 = 0.07906$ and $\beta_1 = 0.90501$ [1]. The sum $B \equiv \alpha_1 + \beta_1$ is then 0.98407, which implies a memory of x_t^2 corresponding to $\tau = 62.3$ trading days. Such a long time memory in the square of returns mimics in an approximate way the power-law correlation of this variable in a finite time window.

Properties of GARCH(1,1) with Gaussian conditional pdf

Another key aspect of the statistical properties of the GARCH(1,1) process is its behavior for different time horizons. For finite variance GARCH(1,1) processes, the central limit theorem applies and one expects that the temporal aggregation of a GARCH(1,1) process progressively implies a decrease in the leptokurtosis of the process. Drost and Nijman [43] carried out a quantitative study of this problem. They were able to show that a 'temporal aggregation' of a GARCH(1,1) process is still a GARCH(1,1) process, but it is characterized by different control parameters. Specifically, when a

Aggregated GARCH(1,1)

GARCH(1,1) x_t is 'aggregated' as

$$S_t^{(m)} = \sum_{i=0}^{m-1} x_{t-i}. \tag{10.21}$$

It can be shown that $S_t^{(m)}$ is also a GARCH(1,1) process characterized by the control parameters [43]

$$\begin{aligned} \alpha_0^{(m)} &= \alpha_0 \frac{1 - B^m}{1 - B} \\ \alpha_1^{(m)} &= B^m - \beta^{(m)} \end{aligned} , \tag{10.22}$$

where $\beta^{(m)} \in (0, 1)$ is the solution of the quadratic equation

$$\frac{\beta^{(m)}}{1 + [\beta^{(m)}]^2} = \frac{\beta_1 B^{m-1}}{1 + \alpha_1^2 [1 - B^{2m-2}]/[1 - B^2] + \beta_1^2 B^{2m-2}}. \tag{10.23}$$

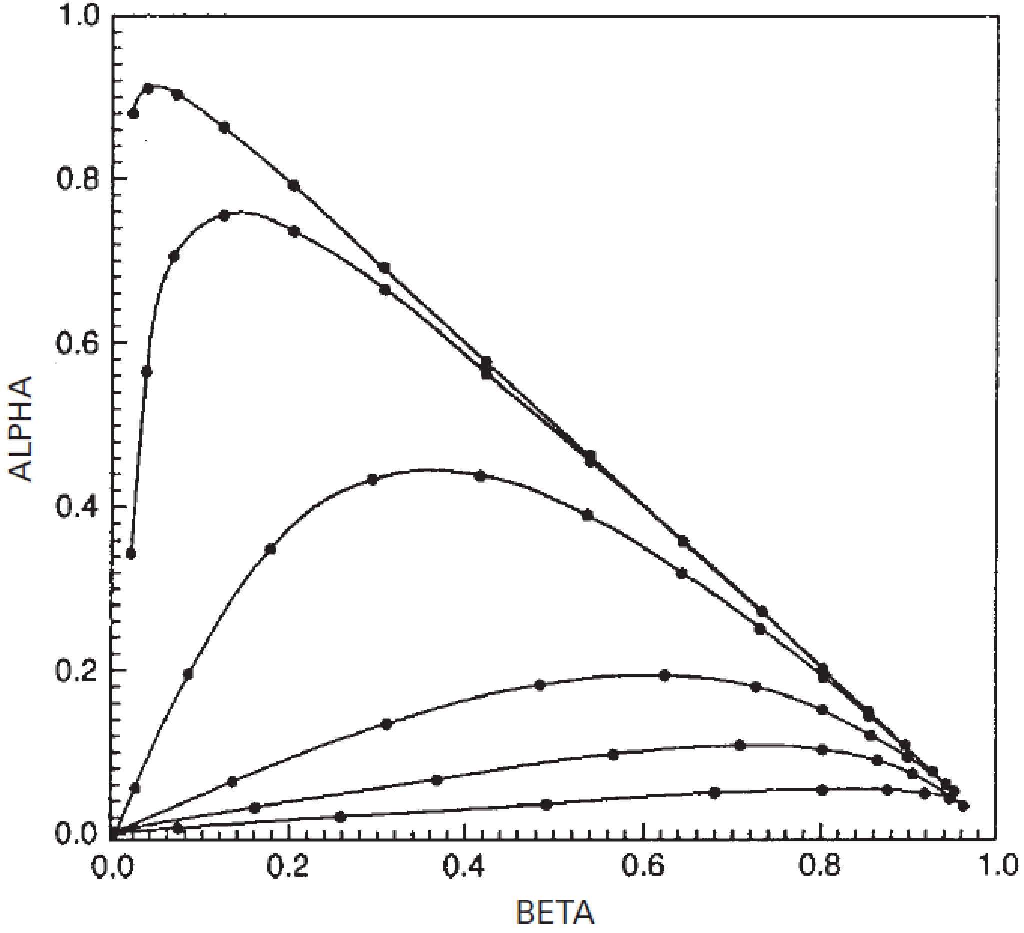


Fig. 10.5. Aggregation of GARCH(1,1). Marks indicate the parameters α_1 and β_1 of a GARCH(1,1) model generated by doubling or halving the sampling interval. The starting GARCH(1,1) processes are characterized by $\beta_1 = 0.8$ and $\alpha_1 = 0.05$, 0.1, 0.15, 0.19, 0.199, and 0.1999 (from bottom to top, respectively).

the attractor for all the GARCH(1,1) processes with finite variance is the process characterized by $\alpha_1^{(m)} = 0$, $\beta^{(m)} = 0$ – namely a Gaussian process.

Properties of GARCH(1,1) with Gaussian conditional pdf

In summary, for any GARCH(1,1) process, temporal aggregation implies that the unconditional pdf of the process presents a degree of leptokurtosis that decreases when the time horizon between the variables increases. Unfortunately, the knowledge of the behavior of $\alpha_1^{(m)}$ and $\beta^{(m)}$ for any value of m is not sufficient to determine the behavior of the probability of return to the origin of a GARCH(1,1) process. We investigate this function numerically in the next section, where we compare empirical findings and GARCH(1,1) simulations.

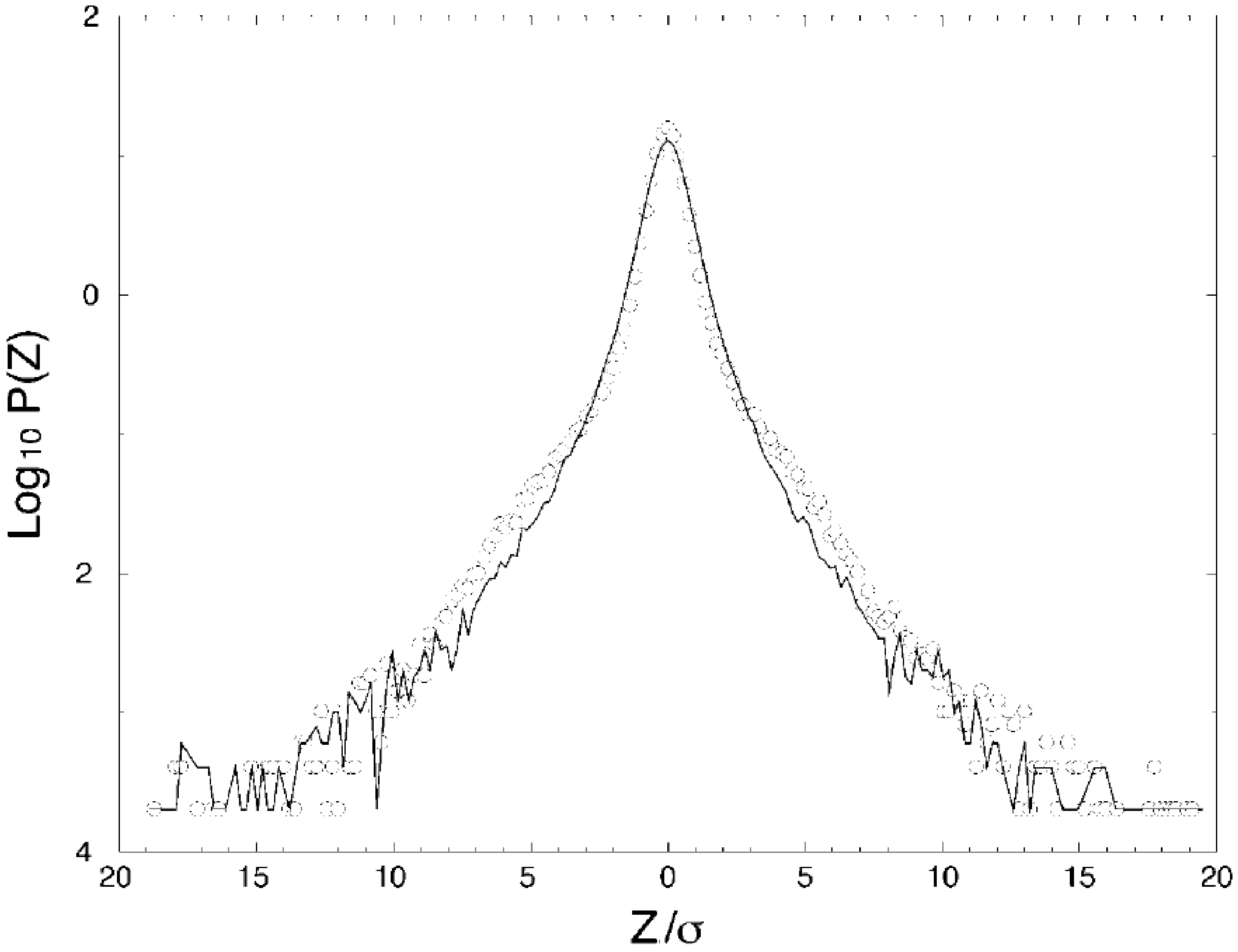


Fig. 10.6. Comparison of the empirical pdf measured from high-frequency S&P 500 data with $\Delta t = 1$ minute with the unconditional pdf of a GARCH (1,1) process characterized by $\alpha_o = 2.30 \times 10^{-5}, \alpha_1 = 0.09105$, and $\beta_1 = 0.9$ (Gaussian conditional probability density). The agreement is good for more than four decades.

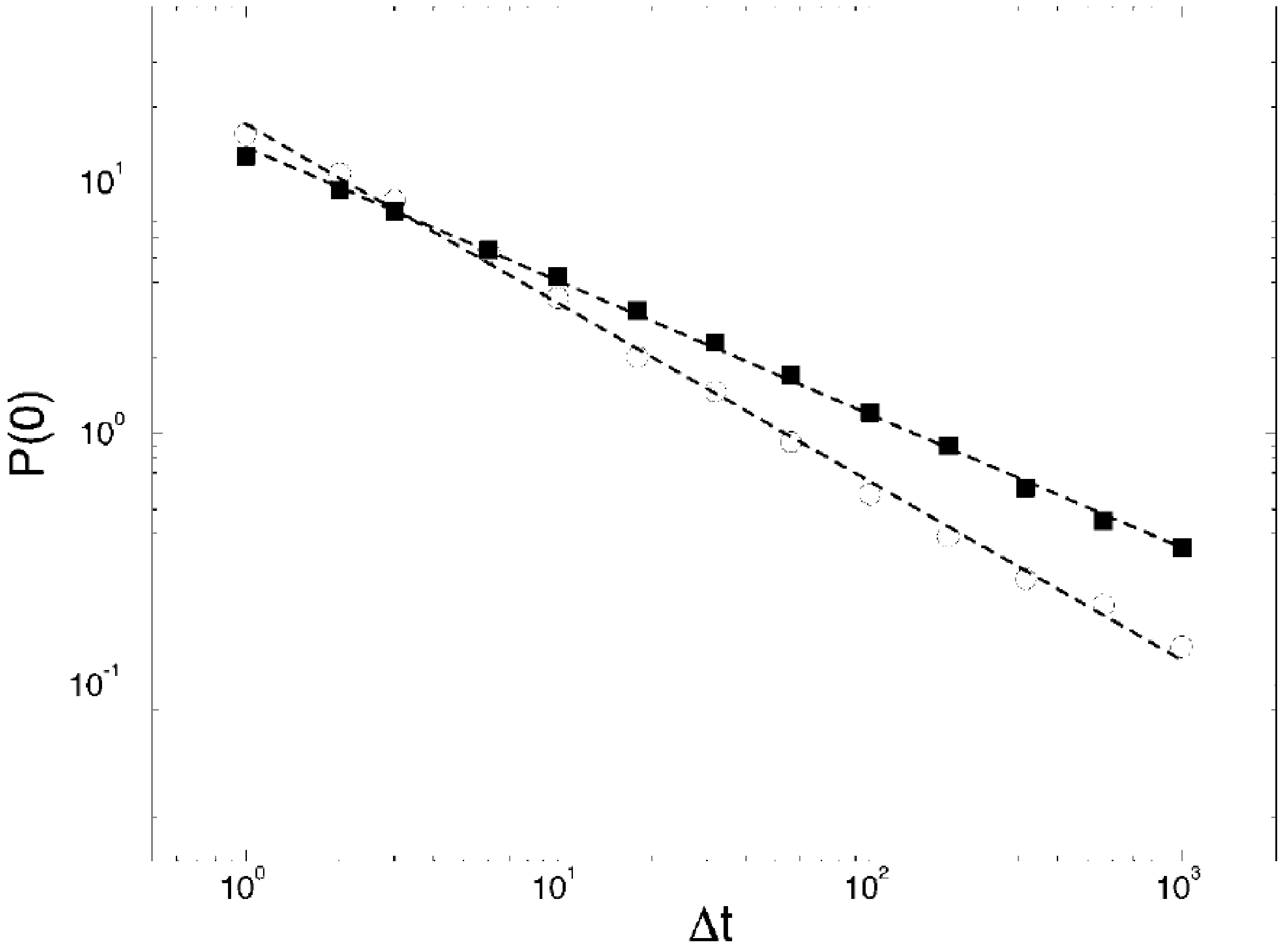


Fig. 10.7. Scaling properties of a GARCH(1,1) stochastic process (black squares), with the same control parameters as in Fig. 10.6. The scaling of the GARCH(1,1) process fails to describe the empirical behavior observed in the S&P 500 high-frequency data (which are also shown for comparison as white circles). Note that the slope 0.53 is extremely close to the Gaussian value of 0.5, indicating that the scaling is close to the scaling of a Gaussian process.

Summary

ARCH and GARCH processes are extremely interesting classes of stochastic processes. They are widely used in finance, and may soon be used in other disciplines. Concerning high-frequency stock market data, ARCH/GARCH processes with Gaussian conditional pdf are able to describe the pdf of price changes at a given time horizon, but fail to describe properly the scaling properties of pdfs at different time horizons.

Open questions concerning this class of stochastic processes include:

(i) What is the form of the asymptotic pdf of the ARCH and GARCH processes characterized by a given conditional probability density function $f_t(x)$?

(ii) What is the nature of the scaling property of the probability of return to the origin as a function of the values of the control parameters and of the shape of the conditional probability density function?

Thank you for your attention!