Value at Risk

Christian Gourieroux[*]

Joann Jasiak[†]

Revised version: December 20, 2002

(First draft: August, 2001)

[*]CREST, CEPREMAP and University of Toronto, *e-mail*: gouriero@ensae.fr.

[†]York University, *e-mail*: jasiakj@yorku.ca

1   Introduction

In this chapter we discuss the management, supervision and measures of extreme and infrequent risks in finance. By extreme risks we mean those that induce very large losses per Dollar invested. When such risks arise infrequently, their occurrence may remain unobserved for a long period of time after which they tend to be forgotten or neglected by investors. A series of recent bank failures, due to mismanaged portfolios invested in corporate loans, real estate or complex derivatives, rose the awareness of bank regulators and prompted the search for new instruments of protection against extreme risks.

By the decision taken at the first meeting of the Governors of Central Banks held in Basle (Switzerland) in 1995, banks were imposed mandatory computation of a risk measure, called **Value at Risk** (VaR), for each line of the balance sheet. Banks were also required to report the VaR, keep it regularly updated and to hold an appropriate amount of capital (the so-called **required capital** (RC)) to hedge against extreme risks.

However, it soon appeared that a number of banks and financial institutions did not possess the necessary data base, know-how and adequate technical equipment to satisfy the above requirements. Consequently, a permanent committee, called the Basle Committee, was established to coordinate the development of a minimal technological stand-by.

The implementation of common guidelines for risk supervision is a very ambitious initiative due to variety of risks that need to be considered. Roughly, the risks can be classified as follows.

i) The **market risk** is due to asset price uncertainty when assets are traded on a competitive market. The market risk is often disregarded when asset prices keep rising for a quite long period of time. In this situation, investors often increase the part of portfolios invested in risky assets and buy options to take advantage of the price increase, but fail to hedge their portfolios against a possible burst of a speculative bubble.

ii) The **credit risk** or **risk of default** is specific to loans when these are evaluated without taking into account the probability of a future payment default.

There exist various types of loans, distinguished with respect to the type of borrower or lender. The borrower can be a consumer, a firm, a County or the Treasury. The loans can be granted directly by a credit institution [over the counter (OTC), also called retail loan] or acquired indirectly by purchasing a bond issued by a firm, a County, or the Treasury.

Let us give some typical examples of neglected default risk. First, consider a credit institution which has established a long term relationship with a customer who in the past took several loans. Suppose that the bank decides to increase substantially the credit line of this customer, given that he had no payment defaults and his rating became high. Such an action may have an adverse

effect, since it requires an increase of the amount of monthly payments which may cause overdebt and default. The following example concerns losses on the bond market due to disregarded risk of default. Such losses might be incurred by investors who believe that a bond with an Aaa rating is riskfree, despite that any company, even a highly rated one, bears the risk of being downrated by Moody's at some point in time.

iii) The **liquidity risk** or **risk of counterparty** exists if it may become difficult to trade quickly a large amount of assets at reasonable prices. Although this risk is often disregarded, in practice it is faced by any company or bank which evaluates at the market price per share [the so-called **marked-to-market** convention] assets listed on its balance sheet. The market price is generally reported for rather small traded volumes and is much higher than the price that would be obtained if the firm had to sell quickly the whole volume of assets that appear on its balance sheet. This explains why, in cases of corporate or bank failures, assets are sold for an amount considerably lower than the one previously computed and reported on the balance sheet.

iv) A number of financial strategies rely on estimation of models for the evolution of asset returns. A typical example is the Black-Scholes model which assumes a geometric brownian motion of asset prices and is used for derivative hedging and option pricing. As noted by several authors, any theoretical model is necessarily misspecified and may provide poor representation of the reality [1]. Moreover, any theoretical model involves some unknown parameters that have to be estimated. The two types of errors due to misspecification and estimation account for the so-called **model risk**. Even if we were very strongly motivated to assess this type of risk, and inspite that it has been recommended to do so by the Basle Committee, such a task is conceptually infeasible since the benchmark, that is the reality, remains unknown.

The primary intention of the Basle Committee is to impose mandatory computation of VaR and of the capital reserve with respect to the four aforementioned notions of risk. In particular, the Basle Committee has launched a long term project of implementing VaR measures for various risks step by step along the lines of the following schedule:

1. VaR for market risk on portfolios of basic liquid assets, such as stocks included in market indexes, Treasury bonds and foreign currencies.

2. VaR for market risk on portfolios that contain also rather liquid derivatives such as options on interest rates, foreign currencies and market indexes.

3. Finally, VaR on portfolios of loans with default risk, called CreditVaR. Two types of assets

---

[1] "I sometimes wonder why people still use the Black-Scholes formula, since it is based on such simple assumptions - unrealistically simple assumptions. Yet that weakness is also its greatest strength. People like the model because they can easily understand its assumptions." (F. Black)

"There are two sources of uncertainty in the prediction made by the trader or the econometrician. The first is parameter estimation, and the second is model error" (Jacquier and Jarrow).

See Merton (1974) for a list of assumptions underlying the Black-Scholes model.

are considered: bonds for which the market prices are available and retail loans for which the bank has insider information about individual credit histories of borrowers.

4. Finally the assessment of model risk is backed-up by the **back testing** procedures for checking the model specification and examining the model-based predictions under extreme scenarios of price evolution (also called **stress testing**).

Our exposition follows the same order. First, we study risk on a portfolio of liquid assets, to which we next add the derivatives, and conclude with a discussion on credit risk. In each section, we present the methods of VaR computation that exist in the literature and point out their advantages and limitations. Among them we indicate those recommended by the Basle Committee. They are not necessarily the most efficient or robust ones, but they always seem to be easy to understand and to implement. Indeed, the objective of the Basle Committee is also to enhance the technological ability of investors [2]. Accordingly, the definitions and computational tools that will be recommended by the Basle Committee in the future are intended to gradually become more and more sophisticated. At the beginning however, the learning process is expected to rely on some straightforward concepts.

The final chapter is devoted to future directions of research and development. We hope that this chapter will give an idea of the content of a similar VaR survey that could be written ten years later.

2 Value at Risk

The aim of this section is to define and compare various notions of Value at Risk for portfolios of assets traded on competitive markets. In particular, it is assumed that assets can be traded at any time, and that the price per share doesn't depend on the traded volume and on whether the transaction is a buy or a sell. Then, the asset price is assumed equal to the ask and bid price, since these are identical. To ensure that this condition is approximately satisfied in practice, the Basle Committee has recommended to use daily data on market closure prices. Indeed, on some stock markets, such as Paris and Toronto, the market closure prices are determined by a market closing auction with a single equilibrium price for each asset.

2.1 Definition

Let us consider a portfolio of n assets, with fixed allocations $a = (a_1, ..., a_n)'$ between t and $t + h$ (say). By allocation we mean quantity and not a monetary value. At date t the investor has

---

[2] In a survey conducted in Australia by Gizycki and Hereford in 1998 a number of portfolios of stocks, bonds, foreign currencies and derivatives of different types were sent to all Australian banks with a request to compute the daily VaR for each of these portfolios according to each bank's own method. Out of all Australian banks only 22 have responded. Out of these, only two banks were able to calculate the VaR for all portfolios.

endowment $W_t(a) = a'p_t$ designated for purchasing this portfolio and an additional reserve amount R_t (say). This reserve amount is supposed to compensate potential adverse changes in the market price (market risk); it has to be put aside and should not be invested on the market. Essentially, the investor selects a reserve amount such that the global position (that is the portfolio value plus the reserve) may incur a loss and become negative, with a predetermined small probability α at date $t + h$. α measures the level of protection. This condition can be written as:

$$P_t[W_{t+h}(a) + R_t < 0] = \alpha, \tag{2.1}$$

where P_t is the conditional distribution of future prices given information I_t used by the investor to predict future prices. Thus $-R_t$ is the α-quantile of the conditional distribution of future portfolio value, called the **profit and loss** [P&L] distribution.

The **required capital** at time t is the sum of the initial endowment plus the reserve. Theoretically, it is equal to the **Value at Risk** denoted by:

$$VaR_t = W_t(a) + R_t, \tag{2.2}$$

and characterized by the condition:

$$P_t[W_{t+h}(a) - W_t(a) + VaR_t < 0] = \alpha. \tag{2.3}$$

It depends on: 1) the information I_t available at time t, 2) the horizon h, 3) the set of assets considered , 4) the portfolio allocation and 5) the loss probability α. These arguments can be introduced explicitly into the VaR formula:

$$VaR_t = VaR(I_t, a, h, \alpha). \tag{2.4}$$

Condition (2.3) is equivalent to:

$$P_t[a'(p_{t+h} - p_t) < -VaR_t] = \alpha. \tag{2.5}$$

Thus the opposite of VaR defined in (2.5) is the α-quantile of the conditional distribution of the change in portfolio value.

Note that

$$
\begin{aligned}
a'(p_{t+h} - p_t) &= \sum_{i=1}^{n} a_i(p_{i,t+h} - p_{i,t}) \\
&= \sum_{i=1}^{n} a_i p_{i,t}[(p_{i,t+h} - p_{i,t})/p_{i,t}]
\end{aligned}
$$

$$= \sum_{i=1}^{n} a_{i,t}^{*} r_{i,t,t+h},$$

where $a_{i,t}^{*}, i = 1, ..., n$ is the portfolio allocation measured in Dollars and $r_{i,t,t+h}, i = 1, ..., n$ are the returns on the asset. Therefore, VaR analysis can be based on returns instead of price changes [3].

VaR can be used for two purposes: 1)to measure risk and 2)to determine the capital reserve. The advantage of VaR is that it provides a better risk measure than the standard volatility while it has the same applications. In particular, it can be used for portfolio management, fund manager auditing, hedging and so on [see e.g. Levy, Sarnat (1972), Arzac, Bawa (1977), Jansen and alii (1998), Foellmer, Leukert (1998)]. In practice, for a more comprehensive understanding of risk, several VaR measures should be computed. This can be done by selecting a set of different risk levels such as $\alpha = 1\%$, 5%...and a set of different horizons, such as h=1, 10, 20 days. As mentioned, Var is used by supervisors to fix the level of reserve. Even though the theoretical value of required capital is identical to the VaR, the Basle Committee has fixed the mandatory required capital RC_t at a different level defined as follows. The banks are required to report daily estimates of the Value at Risk $\widehat{VaR}_t$ at a horizon of 10 business days (i.e. two weeks) and to compute the required capital defined by:

$$RC_t = Max[\widehat{VaR}_t, 3(\text{trigger}/8)\frac{1}{60}\sum_{h=1}^{60}\widehat{VaR}_{t-h}]. \qquad (2.6)$$

This complicated formula is introduced for the following reasons: a) to alleviate the effect of an eventual underestimation of VaR by fixing the multiplicative factor at a value larger than 3, b) to create a positive incentive for a bank to perform the best possible evaluation of risk, by introducing an adjustable trigger value which depends on the ex-post accuracy of the VaR (trigger between 8 and 25), c) to avoid erratic changes in the level of required capital by averaging its lagged values, and d) to allow for quick updating when unexpected markets changes occur.

2.2   Examples

A closed-form expression of VaR can be found for some distributions of price changes (or returns). In this section, we first present the VaR for a conditional Gaussian distribution. Next, we compare the VaR expressions for distributions with different types of tails. Finally, we discuss the dependence of VaR on the holding time, that is the computational horizon h.

i) **The Gaussian Value at Risk**

For convenience we assume a time horizon of length equal to one unit of time $h = 1$. Let us assume that the price changes are conditionally Gaussian with mean $\mu_t = E[p_{t+1} - p_t|I_t]$ and

---

[3]For VaR computation, returns should not be computed as log price differences.

covariance matrix $\Omega_t = V(p_{t+1} - p_t | I_t)$. We get:

$$
\begin{aligned}
& P_t[a'(p_{t+1} - p_t) < -VaR_t] = \alpha \\
\Leftrightarrow \quad & P_t\left[\frac{a'(p_{t+1} - p_t) - a'\mu_t}{[a'\Omega_t a]^{1/2}} < \frac{-VaR_t - a'\mu_t}{[a'\Omega_t a]^{1/2}}\right] = \alpha \\
\Leftrightarrow \quad & -VaR_t - a'\mu_t = \Phi^{-1}(\alpha)(a'\Omega_t a)^{1/2} \\
\Leftrightarrow \quad & VaR_t = -a'\mu_t + \Phi^{-1}(1-\alpha)(a'\Omega_t a)^{1/2},
\end{aligned}
\tag{2.7}
$$

where Φ denotes the cdf of the standard normal distribution. In practice, the predetermined probability of loss is small. Thus, the Value at Risk is an increasing function of volatility of the portfolio value and a decreasing function of a) the expected increment of the portfolio value and b) the loss probability. This approach relies on the two first conditional moments only, and is therefore called the **method of the variance-covariance matrix**.

The required amount of reserve is nonnegative if and only if:

$$
\frac{a'(\mu_t + p_t)}{[a'\Omega_t a]^{1/2}} < \Phi^{-1}(1-\alpha),
$$

that is when the portfolio's Sharpe performance measure is too small. Otherwise the reserve is negative and the investor has a possibility of additional borrowing while satisfying the VaR constraint.

ii)**Comparison of tails**.

Let us consider two portfolios of identical assets, with different allocations a and a^* (say). We assume that their values at date t are equal $a'p_t = a^{*'}p_t$, and we denote by F_t (resp. F_t^*) the conditional cdf of the portfolio value change $y_{t+1} = a'(p_{t+1} - p_t)$ [resp. $y_{t+1}^* = a^{*'}(p_{t+1} - p_t)$]. At risk level α the VaR's for these portfolios are given by:

$$
VaR_t(\alpha, a) = -F_t^{-1}(\alpha), \quad VaR_t(\alpha, a^*) = -(F_t^*)^{-1}(\alpha).
\tag{2.8}
$$

Intuitively, portfolio a^* is more risky than portfolio a if , for any small α, portfolio a^* implies a larger reserve amount than portfolio a. This condition is equivalent to:

$$
\begin{aligned}
& -(F_t^*)^{-1}(\alpha) > -F_t^{-1}(\alpha), \text{ for any small } \alpha, \\
\Leftrightarrow \quad & F_t^*(y) > F_t(y), \text{ for any small } y.
\end{aligned}
$$

Thus, the reserve amount for a^* is larger than for a if and only if the left tail of the conditional distribution of the change of portfolio value is fatter for allocation a^* than for allocation a. If the

portfolio is invested in a single risky asset, the change in the portfolio value is: $a(p_{t+1} - p_t)$. The left tail of the distribution of the change in portfolio value corresponds to the left tail of the asset price change if $a > 0$, and to the right tail otherwise. An investor with a positive amount of this asset is risk averse with respect to a price decrease. An investor is risk averse with respect to the asset price increase if his position is short, that is when he holds a negative quantity of assets. In a multiasset framework the situation is more complicated due to the fact that some asset prices are positively and some are negatively correlated with one another.

It is common to compare extreme risks on two portfolios by considering the limiting left tail behavior of the cdf (when y tends to minus infinity) or the quantile function (when α tends to zero). For example the left tails of one dimensional distributions are often classified in the following way:

i) The distribution F admits a **gaussian left tail** if and only if

$\exists m, \sigma > 0, a > 0 : F(y) \approx a\Phi\left(\frac{y-m}{\sigma}\right)$, when $y \to -\infty$.

ii) The distribution F admits an **exponential left tail** if and only if

$\exists \lambda > 0, a > 0 : F(y) \approx a \exp \lambda y$. λ is called the **tail index** of an exponential tail.

iii) The distribution F admits a **Pareto** (hyperbolic) **left tail** if and only if $\exists \lambda > 0, a > 0 :$ $F(y) \approx a(-y)^{-\lambda}$. λ is called the **tail index** of a Pareto tail.

Asymptotically, Pareto tails are fatter than exponential tails which, in turn, are fatter than gaussian tails. Alternatively, the size of tails can be described in terms of the quantile function or of the VaR. For a distribution with an exponential left tail, the VaR is a logarithmic function of the risk level α: $VaR(\alpha) \approx -\frac{1}{\lambda} \log \frac{\alpha}{a}$ for α small. For a distribution with a Pareto left tail the VaR is a hyperbolic function of α: $VaR(\alpha) \approx (\alpha/a)^{-1/\lambda}$, for α small. Thus the rate at which the tails taper off is directly linked to the rate of increase of the VaR when α tends to zero.

Examples of distributions with exponential left tails are:

the double exponential (Laplace) distribution with cdf:

$$F(y) \quad = \quad \frac{1}{2} \exp \lambda(y - m), \text{ if } y \leq m, \qquad (2.9)$$
$$1 - \frac{1}{2} \exp -\lambda(y - m), \text{ if } y \geq m;$$

the logistic distribution with cdf:

$$F(y) = \left[1 + \exp -\left(\frac{y - m}{\eta}\right)\right]^{-1}.$$

Examples of distributions with Pareto tails are:

the Cauchy distribution with cdf:

$$F(y) = \frac{1}{\pi} \arctan \left(\frac{y-m}{\sigma} \right) + \frac{1}{2};$$

the double Pareto distribution with cdf:

$$\begin{aligned} F(y) &= \frac{1}{2}(\mu - y)^{-\lambda}, \text{ if } y \leq \mu - 1, \\ &\quad 1 - \frac{1}{2}(y - \mu)^{-\lambda}, \text{ if } y \geq \mu - 1. \end{aligned}$$

The analysis of tail sizes and estimation of tail indexes are topics of a large body of literature in statistical theory [see e.g. Embrechts et alii (1999)]. However, any asymptotic comparison of tails has to be interpreted with caution because in finance we are interested in a small, but fixed risk level, such as $\alpha = 1$ %, or 5 %. To illustrate this point, let us consider a logistic distribution with the cdf $F(y) = (1 + \exp y)^{-1}$ and the normal distribution with the same mean and variance as the logistic distribution, that is with mean 0 and variance $\sigma^2 = \pi^2/3$. The VaR computed from the $N(0, \pi^2/3)$ is given by:

$$VaR_N(\alpha) = \pi/\sqrt{3}\Phi^{-1}(1 - \alpha),$$

whereas the VaR computed from the logistic distribution is

$$VaR_L(\alpha) = \ln \left(\frac{1-\alpha}{\alpha} \right).$$

The normal and logistic VaR are ploted in Figure 2.1 as functions of α.

(Insert Figure 2.1 : Comparison of Normal and Logistic VaR)

Once the mean and variance effects are removed, we observe that the VaR curves are rather hard to distinguish except for very small risk levels. In fact the two curves intercept for $\alpha = 4\%$. If $\alpha=5\%$ the gaussian VaR is 2.98 and is larger than the logistic VaR equal to 2.94. If $\alpha=1\%$, the gaussian VaR is 4.22 and is less than the logistic VaR 4.59, although quite close. In summary, we find that: 1)the tail effect on the VaR can be reversed when α increases but remains small; 2) the preliminary adjustments for the mean and variance effects are necessary.

iii)**Term structure of the VaR**

The dependence of VaR on risk level has been discussed in the last section. Now let us focus on the dependence of VaR on the holding time h, henceforth called the horizon. More precisely, we investigate whether VaR can be considered a simple function of h, denoted by $VaR(\alpha, h)$. Some results are easy to derive when the variables y_t, t varying, are i.i.d.. For instance it is easily proved that

$$VaR(\alpha, h) = h^{1/a}b(\alpha), \forall \alpha, \tag{2.10}$$

where a is a scalar and b is a function, if and only if the characteristic function of y_t is of the type:

$$\Psi(u) = E(\exp -iuy_t) = \exp(-c|u|^a).$$

For example, this restriction imposed on the characteristic function is satisfied by a zero mean normal distribution with $a = 2$; thus in this special case the term structure of VaR is such that $VaR_N(\alpha, h) = \sqrt{h}b(\alpha)$. It also holds for the Cauchy distribution with $a = 1$; thus the term structure of the VaR is such that $VaR_C(\alpha, h) = hb^*(\alpha)$. This example shows that the larger the tails, the greater is the impact of the holding period on the value of VaR.

Note that the Basle Committee (1995 p.8) suggested the practice of scaling-up the VaR calculated for a one day holding period by multiplying it by the square root $\sqrt{h} = \sqrt{10} = 3.16$ to obtain the 10 day VaR. This suggestion implicitly assumes independence and normality of price changes.

2.3 Conditional and Marginal Value at Risk

We have seen that VaR depends on the information set used for forecasting the future values of a portfolio. The definition of the information set depends on the approach; It can contain the lagged values of prices of all assets in the portfolio (I_t^1), or contain only the lagged values of the entire portfolio (I_t^2). The information-based approaches lead to computation of the so-called **conditional VaR**. The outcomes of computations conditioned on either (I_t^1) or (I_t^2) are generally not identical. Eventually, it is also possible to disregard all information on past asset prices or portfolio values to get the **marginal VaR** defined by:

$$P[W_{t+h}(a) - W_t(a) + VaR_t < 0] = \alpha, \tag{2.11}$$

where P denotes the marginal probability distribution. When the price changes $(p_t - p_{t-1})$, t varying, are stationary, the marginal VaR is time independent, and the time index can be suppressed. On the contrary the conditional VaR varies in time due to changes in market conditions. For example, if the price changes satisfy a gaussian vector autoregressive (VAR) process of order one: $\Delta p_t = p_t - p_{t-1} = A\Delta p_{t-1} + \epsilon_t$ [4], where $\epsilon_t \sim IIN(0, \Omega)$, A is the $n \times n$ matrix of autoregressive coefficients, Ω the $n \times n$ variance-covariance matrix of the error term, then the conditional distribution of Δp_t given $\Delta p_{t-1}, \Delta p_{t-2}...$ is the gaussian distribution $N(A\Delta p_{t-1}, \Omega)$. Therefore, the conditional VaR is given by: $VaR_t(a, \alpha) = a'A\Delta p_{t-1} + \Phi^{-1}(1 - \alpha)(a'\Omega a)^{1/2}$. On the other

---

[4] The terminology is quite confusing. Note the difference between VaR (Value at Risk) and VAR (Vector Autoregressive process).

hand, we know that the marginal distribution of Δp_t is also gaussian $N(0, \Sigma)$, where $\Sigma = V(\Delta p_t)$ satisfies the equation $\Sigma = A\Sigma A' + \Omega = \sum_{h=0}^{\infty} A^h \Omega A'^h$. Therefore the marginal VaR is given by:

$$VaR(a, \alpha) = \Phi^{-1}(1 - \alpha)(a' \Sigma a)^{1/2}.$$

Since the marginal mean $E\Delta p_t = 0$, and $\Sigma = V(\Delta p_t) >> \Omega$, we find that, on average, the marginal VaR is larger than the conditional VaR.

2.4 Sensitivity of the VaR

Let us consider the VaR at horizon 1 defined by :

$$P_t[a'\Delta p_{t+1} < -VaR_t(a, \alpha)] = \alpha.$$

The Value at Risk depends on portfolio allocation. In practice, a portfolio manager has to update frequently the portfolio frequently, and her major concern is the impact of updating on risk (or on the capital reserve). Thus that manager is more concerned about the dependence of VaR on portfolio allocation than about the value of VaR, iteself.

The analytical expressions of the first and second order derivatives of VaR with respect to portfolio allocation have been derived in Gourieroux, Laurent, Scaillet (2000):

$$\text{i) } \frac{\partial VaR_t(a, \alpha)}{\partial a} = -E_t[\Delta p_{t+1} | a'\Delta p_{t+1} = -VaR_t(a, \alpha)].$$

$$(2.12)$$

$$\text{ii) } \frac{\partial^2 VaR_t(a, \alpha)}{\partial a \partial a'} = \frac{\partial \log g_{a,t}}{\partial z}[-VaR_t(a, \alpha)]V_t[\Delta p_{t+1} | a'\Delta p_{t+1} = -VaR_t(a, \alpha)]$$

$$(2.13)$$

$$+ \left\{ \frac{\partial}{\partial z} V_t[\Delta p_{t+1} | a'\Delta p_{t+1} = z] \right\}_{z = -VaR_t(a, \alpha)} \qquad (2.14)$$

where $g_{a,t}$ denotes the conditional p.d.f. of $a'\Delta p_{t+1}$.

Thus the first and second order derivatives of the VaR can be written in terms of the first and second order conditional moments of price changes in a neighbourhood of the VaR condition : $a'\Delta p_{t+1} = -VaR_t(a, \alpha)$.

The sensitivity of VaR can be examined directly in the gaussian case [see e.g. Garman (1996), (1997)]. Let us denote by μ_t, Ω_t the conditional mean and variance of Δp_{t+1}, respectively. The VaR is given by :

$$VaR_t(a, \alpha) = -a'\mu_t + \Phi^{-1}(1 - \alpha)(a'\Omega_t a)^{1/2}.$$

For example, we get :

$$
\begin{aligned}
\frac{\partial VaR_t(a,\alpha)}{\partial a} &= -\mu_t + \frac{\Omega_t a}{(a'\Omega_t a)^{1/2}}\Phi^{-1}(1-\alpha) \\
&= -\mu_t + \frac{\Omega_t a}{a'\Omega_t a}[VaR_t(a,\alpha) + a'\mu_t] \\
&= -E_t[\Delta p_{t+1}|a'\Delta p_{t+1} = -VaR_t(a,\alpha)].
\end{aligned}
$$

3 Estimation of the Marginal Value at Risk

In this section we discuss estimation of the marginal Value at Risk from historical data on increments of a portfolio value. We denote by $y_t = a'(p_t - p_{t-1})$ the change of portfolio value and by VaR the marginal VaR at risk level α. Thus, for notational convenience we don't introduce explicitly the dependence on the portfolio allocation a, the risk level α, and the horizon h set equal to one.

The marginal VaR is given by:

$$
P[y_t < -VaR] = \alpha. \tag{3.1}
$$

It means that the opposite of the marginal VaR is equal to the α-quantile of the marginal distribution of y_t that can be characterized in two different ways. First, for the marginal cumulative distribution function of y_t, denoted by F, the VaR is defined by:

$$
F(-VaR) = \alpha \Leftrightarrow VaR = -F^{-1}(\alpha). \tag{3.2}
$$

Alternatively, the VaR can be defined as a solution to the following minimization:

$$
-VaR = \underset{\theta}{Argmin}\, E[\alpha(y_t - \theta)^+ + (1-\alpha)(y_t - \theta)^-], \tag{3.3}
$$

where $(y_t - \theta)^+ = Max(y - \theta, 0), (y_t - \theta)^- = Max(\theta - y, 0)$. It is easy to show that the first order condition of the minimization leads to equation (3.1).

In the sequel, these two characterizations are used to define various VaR estimators in the parametric, semi-parametric and nonparametric frameworks. The properties of estimators will be discussed for i.i.d. observations $y_t, t = 1, ..., T$, for which the marginal VaR coincides with the conditional one (see section 2.2). The last section is concerned with the properties of these estimators, when the i.i.d. asssumption is violated.

3.1 Historical Simulation

A natural estimation method that does not rely on any assumption about the marginal distribution of y_t, consists of considering the empirical counterparts of expressions (3.3) and (3.2). We get:

$$\widehat{VaR} = \underset{\theta}{Argmin} \frac{1}{T} \sum_{t=1}^{T} [\alpha(y_t - \theta)^+ + (1 - \alpha)(y_t - \theta)^-], \qquad (3.4)$$

or

$$\widehat{VaR} = \underset{\theta}{Argmin}[\hat{F}(-\theta) - \alpha]^2, \qquad (3.5)$$

where $\hat{F}(y) = \frac{1}{T} \sum_{t=1}^{T} 1_{y_t < y}$ denotes the empirical c.d.f., and $1_{y_t < y} = 1$ when $y_t < y$ and 0 otherwise, is the indicator function.

Thus the theoretical α- quantile is approximated by an empirical α- quantile. The minimizations in (3.4) and (3.5) provide two equivalent methods for computing an empirical quantile from a sample. However, it is important to note that the solutions are generally not unique. More precisely, since the empirical distribution is discrete, the empirical c.d.f. is not continuous, and we obtain an interval of solutions, called the empirical α- **quantile interval**.

In practice, a solution is easily obtained in the following way. Let us assume that $T = 200$ and $\alpha = 1\%$. The observations $y_t, t = 1, .., 200$ can be ranked in ascending order

$y_{(1)} = Min_t y_t < y_{(2)}...... < y_{(T)} = Max_t y_t$. Then, the 1%-quantile interval is $[y_{(2)}, y_{(3)}]$. Its lower bound corresponds to the second smallest observation, since $2/200 = 1/100$, and the upper bound is equal to the next observation.

The asymptotic properties of the empirical quantile were derived by Basset, Koenker (1978) [see also Gourieroux, Monfort (1998), section 8.5] for i.i.d. data.

When T tends to infinity and the risk level α is fixed:

i) the length of the empirical α- quantile interval tends to zero;

ii) any value in the empirical α- quantile interval is a consistent and asymptotically normal estimator:

$$\sqrt{T}(\widehat{VaR} - VaR) \sim N\left[VaR, \frac{F(-VaR)[1 - F(-VaR)]}{f^2(-VaR)}\right], \qquad (3.6)$$

where f denotes the marginal p.d.f. of y_t.

For small α, the asymptotic variance of the VaR estimate depends on the risk level and left tail pattern of the marginal distribution. The estimator is less accurate when α diminishes, or the left tail increases. To illustrate this point, let us consider a distribution with hyperbolic tails [called Pareto tails]. If $F(y) \sim c(-y)^{-\beta}$, for small α the asymptotic variance:

$$V_{as}[\sqrt{T}(\widehat{VaR} - VaR)] \approx \frac{F(-VaR)}{f^2(-VaR)} \approx \frac{1}{\beta^2 c}(-VaR)^{\beta+2},$$

is an increasing function of the tail parameter β.

The asymptotic results given above are intended to help to understand the accuracy of the empirical quantile. However, they are not valid when T tends to infinity and α tends to zero,

which is the situation encountered in finance when, for instance, T is large and α is small. To give some insights on this situation, let us consider a sample of $T = 100$ for which an empirical 1%-quantile is $Z = y_{(1)} = Min_t\, y_t$. The finite sample distribution of $y_{(1)}$ is easy to find. Indeed, we have : $P[Z > z] = P[Min_t\, y_t > z] = P[y_t > z, t = 1, ...100] = [1 - F(z)]^{100}$. Then the median $z_{0.5}$ of the 1%-quantile Z is given by:

$$[1 - F(z_{0.5})]^{100} \quad \Leftrightarrow \quad z_{0.5} = F^{-1}[1 - (0.5)^{1/100}] \tag{3.7}$$

$$\Leftrightarrow 1 - F(z_{0.5}) = (0.5)^{1/100} \quad \Leftrightarrow \quad z_{0.5} \approx F^{-1}\left[-\frac{ln0.5}{100}\right] = F^{-1}\left[\frac{0.7}{100}\right]. \tag{3.8}$$

For instance, when the marginal distribution is a Cauchy distribution with c.d.f. $F(y) = \frac{1}{\pi}Arctg y + \frac{1}{2}$, the theoretical 1%-quantile is equal to -31.8, whereas the distribution of the empirical quantile Z admits the median -45.5 and a 90% prediction interval [-636.6, -10.5]. It is clear that the finite sample distribution of the empirical quantile is skewed and its median is far from the true value. Therefore, the result on asymptotic normality does not apply to this case.

The finite sample properties of the empirical quantile estimators can be illustrated by Monte-Carlo experiments. We consider $T = 200$ observations and risk levels $\alpha = 1\%$ and 5%. Figure 3.1 shows the finite sample distribution of $y_{(1)}, y_{(2)}, y_{(3)}$ associated with $\alpha = 1\%$ when the true distribution F is N(0,1), two-sided exponential and Cauchy, respectively. The true values of the 1%-quantile are -2.33, -3.91, -31.82, respectively.

(Insert Figure 3.1: Finite Sample Distributions of the 1%-empirical quantile)

Typically, for the gaussian and two-sided exponential distributions we observe that the empirical quantile $y_{(3)}$ is less biased than the empirical quantile $y_{(2)}$. Moreover, for Cauchy data, the left tail of the distribution of the empirical quantile is very heavy and this estimator has poor accuracy.

Figure 3.2 shows similar results for $y_{(9)}, y_{(10)}, y_{(11)}$ and the level $\alpha = 5\%$. The true values of the 5%-quantile are -1.64, -2.30, -6.31, respectively.

(Insert Figure 3.2: Finite Sample Distributions of the 5%-empirical quantile)

To conclude, estimation of the theoretical α-quantile by its empirical counterpart is an appealing approach to VaR estimation, but leads to very inaccurate and unstable estimates, especially for small risk levels.

Finally, let us point out another limitation of the empirical quantile approach. In practice, it is common to compute this estimate for different risk levels and different portfolio allocations, from the same set of basic asset price changes. While the true underlying VaR is a continuously differentiable function of the arguments α and a , this property is not satisfied by the associated estimate $\widehat{VaR}$. Due to discreteness, a small change in portfolio allocation can cause a jump in the value of the estimate, whereas the true underlying VaR is slowly varying. This drawback can

be partly eliminated by smoothing the expression that defines the estimator. For example, let us consider a kernel K and a bandwidth h; an estimator that is smoothed with respect to α and a is the solution of:

$$\frac{1}{T}\sum_{t=1}^{T} K\left[\frac{-a'(p_t - p_{t-1}) - \widehat{VaR}(a,\alpha)}{h}\right] = \alpha, \tag{3.9}$$

where the dependence on α and a is made explicit. It can be shown that the asymptotic properties of this estimator are identical to the asymptotic properties of the empirical α-quantile, when α is fixed, T tends to infinity and the bandwidth h tends to zero at an appropriate rate [Falk (1985), Horowitz (1992)]. In fact, our discussion shows that one should be more interested in estimating the functional parameter $(a, \alpha) \to VaR(a, \alpha)$ rather than in finding a specific value of this function. In the functional approach it seems natural to impose on the functional estimator $(a, \alpha) \to \widehat{VaR}(a, \alpha)$ the regularity properties satisfied by the underlying theoretical VaR function.

3.2   Parametric Methods

Let us assume that the price changes $\Delta p_t = p_t - p_{t-1}$ admit a distribution that belongs to a parametric family with pdf g_θ and parameter θ. The parameter θ can be estimated by the maximum likelihood yielding:

$$\hat{\theta}_T = \underset{\theta}{Argmax}\sum_{t=1}^{T} \ln g_\theta(\Delta p_t).$$

Then the VaR can be approximated by

$$\widehat{VaR} = -F_{a,\hat{\theta}_T}^{-1}(\alpha),$$

where $F_{a,\theta}$ is the cdf of $\Delta W_t(a) = a'\Delta p_t$, when Δp_t follows g_θ.

In particular, for gaussian price changes $\Delta p_t \sim N(\mu, \Omega)$, the VaR is simply estimated by:

$$\widehat{VaR} = -a'\hat{\mu} + \Phi^{-1}(1 - \alpha)(a'\hat{\Omega}a)^{1/2}, \tag{3.10}$$

where $\hat{\mu}$ and $\hat{\Omega}$ are the sample mean and covariance matrix computed from the observations $\Delta p_1, ..., \Delta p_T$. Since in the gaussian case the finite sample distribution of $(a'\hat{\mu}, a'\hat{\Omega}a)$ is known, we can find the distribution of the VaR estimator as well as the corresponding confidence level. Even though this feature is often neglected by practitioners, one should keep in mind that the estimated measure of risk is random and therefore risky.

3.3 Semiparametric Analysis

When the loss probability α is small, the empirical quantile is estimated from a limited number of extreme observations. Therefore the estimated VaR may not be accurate. Alternatively, in the parametric approach the basic model can be misspecified, which would cause a bias in VaR estimation. To circumvent both difficulties we can estimate empirically the quantiles for some rather large values of α (say) and deduce the VaR of interest from a parametric model of the tail. In practice, the parametric model can be based on a Pareto distribution, or on an exponential distribution, or else on a mixture of two normal distributions [Longerstay (1996), Venkataraman (1997)]. This approach is called the **model building** method. Let us assume, for instance, a Pareto-type model for the left tail where we have approximately :

$$F(y) \sim c(-y)^{-\beta}.$$

where c and β are positive. In the following paragraphs we describe two alternative methods for estimation of c and β for the tail and we derive the associated VaR.

  i) **Estimation by empirical quantiles**

  Let us consider two rather high values of risk level, such as $\alpha_0 = 10\%, \alpha_1 = 5\%$. For such risk levels, the empirical quantiles, denoted $\widehat{VaR}_e(\alpha_0)$ and $\widehat{VaR}_e(\alpha_1)$ are quite accurate and we get the approximate moment conditions:

$$\begin{cases} \alpha_0 & \sim & c[\widehat{VaR}_e(\alpha_0)]^{-\beta} \\ \alpha_1 & \sim & c[\widehat{VaR}_e(\alpha_1)]^{-\beta} \end{cases}$$

  By solving the system of equations with respect to c and β, we obtain consistent estimators of the parameters for the tail:

$$\begin{cases} \alpha_0 & \sim & \hat{c}[\widehat{VaR}_e(\alpha_0)]^{-\hat{\beta}} \\ \alpha_1 & \sim & \hat{c}[\widehat{VaR}_e(\alpha_1)]^{-\hat{\beta}} \end{cases}$$

Then the VaR at any small risk level α ($\alpha = 1\%$, say) can be estimated by:

$$\alpha = \hat{c}[\widehat{VaR}_e(\alpha_1)]^{-\hat{\beta}}. \tag{3.11}$$

  The last two systems of equations are linear in $\hat{\beta}$ and $\log \hat{c}$.

$$\begin{cases} \log \alpha_0 & = & \log \hat{c} - \hat{\beta} \log \widehat{VaR}_e(\alpha_0), \\ \log \alpha_1 & = & \log \hat{c} - \hat{\beta} \log \widehat{VaR}_e(\alpha_1) \\ \log \alpha & = & \log \hat{c} - \hat{\beta} \log \widehat{VaR}(\alpha) \end{cases}$$

Therefore the estimated $\widehat{VaR}(\alpha)$ is related to the initially computed empirical quantiles $\widehat{VaR}_e(\alpha_0)$ and $\widehat{VaR}_e(\alpha_1)$ by:

$$\det \begin{pmatrix} \log \alpha_0 & 1 & \log \widehat{VaR}_e(\alpha_0) \\ \log \alpha_1 & 1 & \log \widehat{VaR}_e(\alpha_1) \\ \log \alpha & 1 & \log \widehat{VaR}(\alpha) \end{pmatrix} = 0,$$

or equivalently:

$$\widehat{VaR}(\alpha) = [\widehat{VaR}_e(\alpha_1)]^A [\widehat{VaR}_e(\alpha_0)]^B, \tag{3.12}$$

where $A = \frac{\log \alpha - \log \alpha_0}{\log \alpha_1 - \log \alpha_0}$ and $B = \frac{\log \alpha_1 - \log \alpha}{\log \alpha_1 - \log \alpha_0}$.

This formula allows extrapolation of the values $\widehat{VaR}(\alpha)$, for any small α, from two benchmark empirical quantiles. We find that the Pareto-type model of tail implies a geometric extrapolation formula.

ii) **The use of Hill estimator**

When a distribution has exactly a Pareto left tail, then

$$F(y) = c(-y)^{-\beta}, \text{ for } y < \underline{y}, \tag{3.13}$$

where $\underline{y}$ is a given threshold. Then, we can apply the maximum likelihood estimation to the right truncated observations y_t such that $y_t < \underline{y}$ [5]. The truncated Pareto distribution admits the cdf:

$$F_{\underline{y}}(y) = \frac{F(y)}{F(\underline{y})} = \left(\frac{y}{\underline{y}}\right)^{-\beta},$$

which involves only the tail parameter β. The truncated ML estimator of β is:

$$\hat{\beta} = \underset{\beta}{Argmax} \sum_{t=1}^{T} 1_{y_t < \underline{y}} [\log \beta - (\beta + 1)(\log y_t - \log \underline{y})].$$

The first order condition yields:

$$1/\hat{\beta} = \sum_{t=1}^{T} 1_{y_t < \underline{y}} (\log y_t - \log \underline{y}). \tag{3.14}$$

This estimator was first proposed by Hill (1975).

The Hill estimator can be used jointly with an empirical quantile estimator to recover the VaR along the following lines [see e.g. Danielsson, DeVries (1997), (1998)]. Let the risk level α_0 be fixed at a rather high value ($\alpha_0 = 10\%$, say). We consider the empirical quantile $\widehat{VaR}_e(\alpha_0)$ and the Hill estimator associated with $\underline{y} = -\widehat{VaR}_e(\alpha_0)$. The relation:

$$\alpha_0 \approx c[-\widehat{VaR}_e(\alpha_0)]^{-\hat{\beta}},$$

[5]The use of the maximum likelihood approach restricted to tails has been recommended by Embrecht, Resnick, Samorodnitsky (1998) to estimate the VaR.

is used to get a consistent estimator of c:

$$\hat{c} = \alpha_0[-\widehat{VaR}_e(\alpha_0)]^{\hat{\beta}}.$$

Then the VaR associated with any small risk level α is approximated by $\widehat{VaR}(\alpha)$ such that:

$$\alpha = \hat{c}[-\widehat{VaR}(\alpha)]^{-\hat{\beta}}$$
$$\Leftrightarrow \quad \widehat{VaR}(\alpha) = \widehat{VaR}_e(\alpha_0) \left(\frac{\alpha_0}{\alpha}\right)^{1/\hat{\beta}}.$$

Thus the empirical quantile $\widehat{VaR}_e(\alpha_0)$ is multiplied by a scale factor which is a power function of α_0/α. When β increases the tail of the Pareto distribution decreases and so does the VaR.

3.4 The i.i.d. Assumption

Since the marginal and conditional VaR are equal for i.i.d. price changes, it seemed to us natural to present the statistical properties of VaR estimators in the i.i.d. framework. However the i.i.d. assumption is not satisfied in practice. This is known due to empirical evidence on serial correlation of price changes, conditional heteroscedasticity and volatility persistence. Moreover, for theoretical reasons the i.i.d. assumption cannot be satisfied both by price changes Δp_t and by returns $r_t = \Delta p_t/p_{t-1}$. But, even if it were satisfied for price changes (or returns) at horizon 1, it wouldn't hold for price changes (or returns) at any horizon h: $\Delta^h p_t = p_t - p_{t-h}$ since the intervals $\{t-h, t\}$ and $\{t-1-h, t-1\}$ overlap, which implies correlation between $\Delta^h p_{t-1}$ and $\Delta^h p_t$.

The i.i.d. assumption can be violated in two different ways. First, the conditional and marginal distributions may be identical, but time dependent. Then the process is nonstationary. If the variation in time is smooth, then the distributions at close dates look similar. Second, the price changes may be stationary but feature serial dependence. In that case the marginal and conditional distributions are different. Still, the conditional distributions at different dates can resemble one another. This can happen when the price histories at those dates don't differ too much.

If the i.i.d. assumption is imposed for computation of the marginal VaR, then any departure from the i.i.d. framework leads to misspecification and can have serious consequences. In the case of a nonstationary process without serial dependence, misspecification is due to replacing the true time varying marginal VaR by a constant, time invariant VaR obtained from such calculations. Then the constant estimators are not consistent.In the case of a stationary process with serial dependence, the estimators are still consistent, but their variance is different from the variance assessed under the i.i.d. assumption.

The guidelines of the Basle Committee are set out as if the marginal distributions of price changes (returns) were varying smoothly in time. Therefore to eliminate the bias in VaR estimation,

the Committee has recommended to estimate the marginal VaR by rolling. More precisely it is proposed to select a window of T_0 observations (with a minimum of one year, i.e. $T_0 > 200$). At date t, the estimation is performed from the T_0 most recent observations: $y_t, y_{t-1}, ..., y_{t-T_0+1}$. At time $t+1$ the newly arrived observation y_{t+1} is added to the sample, while the oldest one y_{t-T_0+1} is deleted. The approach by rolling provides time varying estimates of VaR. It can be improved by introducing exponentially weighted moving averages [Phelan (1995)].

To illustrate the application of rolling estimators, we perform a Monte-Carlo experiment, in which the changes of portfolio value are i.i.d. and the marginal VaR is estimated by rolling from a window of 200 observations. Two estimators were considered: the rolling empirical quantile and the rolling gaussian VaR at 1%. Moreover we considered three sets of i.i.d. simulations with gaussian, double exponential and Cauchy distributions, respectively. The simulation results are displayed in Figures (3.3) - (3.5).

(Insert Figure 3.3: i.i.d. gaussian price changes)

(Insert Figure 3.4: i.i.d. double exponential price changes)

(Insert Figure 3.5: i.i.d. Cauchy price changes)

We observe that the evolution of the rolling empirical quantile in time is a stepwise function. This is easy to explain. Let us consider for example the 1% empirical quantile. A change in the value of the empirical quantile between date t and $t+1$ occurs only if the new and the deleted observations are neither greater, nor less than the value of the quantile estimated at t. This occurs with probability $1 - (1\%)^2 - (99\%)^2 = 0.0198$. Therefore the value of the empirical quantile remains constant for a random time, which has a geometric distribution with parameter equal to $(1\%)^2 + (99\%)^2 = 0.980$. The corresponding average duration is about 50 [6].

It can also be noted that the outcomes of the rolling empirical quantile and gaussian VaR estimations are close for gaussian price changes, whereas the gaussian VaR is less than the empirical quantile for exponentially distributed price changes. This is due to underestimation of the exponential tail when a gaussian VaR formula is used. However, we can also observe the converse effect while comparing the gausssian data to the Cauchy distributed ones. Indeed the first and second order moments of the Cauchy distribution do not exist. Moreover, it is known that the sample mean and variance of Cauchy variables don't converge and feature fat tail asymptotic distributions [7]. Finally, the rolling computation induces a spurious trend effect in the evolution of VaR.

---

[6]To avoid the stepwise effect, it has been proposed by Hull to compute a weighted empirical quantile, solution of $\widehat{VaR}_t = Argmin_\theta \sum_{\tau=t-T_0+1}^{t} \rho^{t-\tau} [\alpha(y_\tau - \theta)^+ + (1-\alpha)(y_\tau - \theta)^-]$.

[7]For example, the empirical average of independent Cauchy variables also admits a Cauchy distribution.

4   Estimation of the Conditional Value at Risk

As mentioned in section 2.3, two types of conditional Value at Risk can be considered according to the selected information set which may contain the lagged values of price changes for each asset in the portfolio (I_t^1), or the lagged portfolio values only (I_t^2). The estimation methods can be parametric, semi-parametric or nonparametric.

4.1   Conditionally Heteroscedastic Autoregressive Models

A common approach to modelling the price change dynamics (or returns) is based on conditionally heteroscedastic autoregressive models.

When the conditioning information set is (I_t^1), the specification is multivariate:

$$\Delta p_t = \mu(\underline{\Delta p_{t-1}}) + B(\underline{\Delta p_{t-1}})u_t, \tag{4.1}$$

where $(\underline{\Delta p_{t-1}}) = (\Delta p_{t-1}, \Delta p_{t-2}, ...)$ denotes the set of lagged values of the price changes, μ is a n-dimensional vector of conditional location parameters, B is a $n \times n$ matrix of conditional scale parameters, and u_t, t varying, is a sequence of i.i.d. random vectors, with common distribution with pdf g.

For the information set (I_t^2), the specification is univariate:

$$y_t = m(\underline{y_{t-1}}) + \sigma(\underline{y_{t-1}})v_t, \tag{4.2}$$

where m and σ are scalar functions and v_t, t varying, is a sequence of i.i.d. variables. Special cases of this specification are the ARCH and GARCH models, such as the ARCH(1) model [see Engle (1982)]:

$$y_t = (\theta_0 + \theta_1 y_{t-1})^2 v_t,$$

and the IGARCH model:

$$y_t = [(1 - \theta) \sum_{j=1}^{\infty} \theta^{j-1} y_{t-j}^2]^{1/2} v_t.$$

The link between both specifications has to be examined with caution. For simple illustration, let us consider the linear vector autoregressive process of order one of asset price changes:

$$\Delta p_t = A\Delta p_{t-1} + \Omega^{1/2}u_t,$$

where u_t is $IIN(0, Id)$, and Id denotes the identity matrix. We know that the conditional distribution of $y_t = \Delta W_t(a) = a'\Delta p_t$, given I_t^1, is gaussian $N[a'A\Delta p_{t-1}, a'\Omega a]$, whereas the conditional

distribution of $\Delta W_t(a)$ given the lagged portfolio values (I_t^2) is of the type $N[\sum_{j=1}^{\infty} \alpha_j a' \Delta p_{t-j}, \sigma^2]$. Thus the specification based on (I_t^2) leads to a univariate model $y_t = m(\underline{y_{t-1}}) + \sigma v_t$, $v_t \sim IIN(0,1)$ with an infinite autoregressive lag which differs from the univariate specification for y_t conditional on the full information set (I_t^1). The situation is much more complex in nonlinear and nongaussian frameworks. Indeed, if the location and scale parameters in (4.1) are nonlinear and if u_t is nongaussian, it is always possible to compute numerically the univariate conditional distribution of $y_t = \Delta W_t(a)$ given its own past, but this conditional distribution may be incompatible with any univariate nonlinear model such as (4.2).

Until now we assumed that the allocation vector a is fixed. It can also be shown that a set of nonlinear autoregressive models (4.2) written for several portfolios that differ with respect to allocations, may be incompatible. For example, let us consider a gaussian ARCH(1) model:

$$\Delta W_t(a) = [\theta_0(a) + \theta_1(a)\Delta W_{t-1}(a)^2]^{1/2} v_t(a), \ \forall a,$$

where $v_t(a)$ is $IIN(0,1)$. The models for $\Delta W_t(a), a$ varying, are compatible only if $\theta_1(a) = 0, \forall a$, that is in the absence of conditional heteroscedasticity. This lack of coherency needs to be emphasized because of a practice initiated by J.P. Morgan (1995) and adopted by the Basle Committee too. The Committee has suggested to use systematically (that is for any set of assets and any portfolio allocation) the gaussian IGARCH model defined by:

$$\Delta W_t(a) = \sigma_t v_t,$$

where

$$
\begin{aligned}
\sigma_t^2 &= \theta \sigma_{t-1}^2 + (1-\theta)\Delta W_{t-1}(a)^2 \\
&= (1-\theta)\sum_{j=1}^{\infty} \theta^{j-1}\Delta W_{t-j}(a)^2 \\
&= (1-\theta)\sum_{j=1}^{\infty} \theta^{j-1} y_{t-j}^2,
\end{aligned}
$$

with $\theta = 0.95$. This model can be valid for some allocations, but it cannot hold for all a's simultaneously.

i) Estimation under the information on portfolio value

The estimation of the VaR conditional on I_t^2 is quite straightforward when the location and scale functions m and σ are parametrized:

$$y_t = m(\underline{y_{t-1}}; \theta) + \sigma(\underline{y_{t-1}}; \theta)v_t. \tag{4.3}$$

The parameter θ can be consistently estimated by the quasi (pseudo) maximum likelihood. The estimator $\hat{\theta}_T$ is given by:

$$\hat{\theta}_T = \underset{\theta}{Argmax} -\frac{1}{2}\sum_{t=1}^{T}\log\sigma^2(\underline{y_{t-1}};\theta) - \frac{1}{2}\sum_{t=1}^{T}\frac{(y_t - m(\underline{y_{t-1}};\theta))^2}{\sigma^2(\underline{y_{t-1}};\theta)}. \quad (4.4)$$

Given this estimate, we can find the approximations of the conditional drift and volatility: $\hat{m}_t = m(\underline{y_{t-1}};\hat{\theta}_T), \hat{\sigma}_t = \sigma((\underline{y_{t-1}};\hat{\theta}_T)$, and the standardized residuals:

$$\hat{v}_t = \sum_{t=1}^{T}\frac{y_t - m(\underline{y_{t-1}};\hat{\theta}_T)}{\sigma(\underline{y_{t-1}};\hat{\theta}_T)}, \quad (4.5)$$

that provide approximations of the true errors v_t. Then , computation of the conditional VaR at horizon 1 can be accomplished by calculating a marginal VaR at horizon 1 from the i.i.d. error terms. Indeed, at horizon 1, we get:

$$P_t[y_{t+1} < VaR_t(a,\alpha,1)] = \alpha$$
$$\Leftrightarrow \quad P_t[m(\underline{y_{t-1}};\theta) + \sigma(\underline{y_{t-1}};\theta)v_t < -VaR_t(a,\alpha,1)] = \alpha$$
$$\Leftrightarrow \quad P_t[v_t < -\frac{VaR_t(a,\alpha,1) - m(\underline{y_{t-1}};\theta)}{\sigma(\underline{y_{t-1}};\theta)}] = \alpha.$$

Therefore $-[VaR_t(a,\alpha,1) - m(\underline{y_{t-1}};\theta)]/\sigma(\underline{y_{t-1}};\theta)$ is the α-quantile of the distribution of the standardized errors v_t .

Since the errors v_t, t varying, (resp. the residuals) are i.i.d. (resp. approximately i.i.d.), we can estimate the α-quantile of the v distribution. For example, a parametric method can be used. When the distribution of the error is assumed gaussian, the conditional VaR is estimated by:

$$\widehat{VaR}_t(a,\alpha,1) = -\hat{m}_t + \hat{\sigma}_t\Phi^{-1}(1-\alpha).$$

However, it has been observed that the residuals often have fat tail distributions. Therefore, several authors proposed parametric models with t-student or α-stable distributions of the error term.

Alternatively, a nonparametric approach can be followed, which relies on the empirical distribution of the residuals $\hat{v}_1, ..., \hat{v}_T$. For the α-quantile denoted by $\hat{Q}(\alpha)$ and computed from the residuals, the VaR estimate is:

$$\widehat{VaR}_t(a,\alpha,1) = -m(\underline{y_{t-1}};\hat{\theta}_T) - \sigma(\underline{y_{t-1}};\hat{\theta})\hat{Q}(\alpha). \quad (4.6)$$

An analytical formula of the VaR estimator exists only for horizon 1. Typically, under a parametric approach where (y_t) satisfies a conditionally gaussian ARCH model, the conditional distribution at horizon 2 is no longer gaussian and has a complicated expression. Hence, for any horizon

larger than 1, the VaR estimate has to be derived by Monte-Carlo methods, and the nonlinear conditional autoregressive model is adequate for conditional simulation at any horizon. For ease of exposition, let us assume the autoregressive order equal to one and consider the nonparametric approach. We get:

$$y_{t+1} = m(y_t; \theta) + \sigma(y_t; \theta)v_{t+1},$$

and

$$y_{t+2} = m(y_{t+1}; \theta) + \sigma(y_{t+1}; \theta)v_{t+2}.$$

Then y_{t+2} can be approximately simulated from the conditional distribution of y_{t+2} given y_t along the following lines [8]. For v_{t+1}^s, v_{t+2}^s two independent drawings in the sample distribution of residuals, a simulated value of y_{t+2} given the currently observed $y_t = y$ is:

$$y_{t+2}^s = m(y_{t+1}^s; \hat{\theta}_T) + \sigma(y_{t+1}^s; \hat{\theta}_T)v_{t+2}^s,$$

where

$$y_{t+1}^s = m(y_t^s; \hat{\theta}_T) + \sigma(y_t^s; \hat{\theta}_T)v_{t+1}^s.$$

By replicating this procedure S times, we obtain a set of values $y_{t+2}^s, s = 1, ..., S$ approximately independently drawn in the conditional distribution of y_{t+2} given the currently observed $y_t = y$. An estimator of VaR$(a, \alpha, 2)$ is the empirical quantile computed from $y_{t+2}^s, s = 1, ..., S$. Note that the chosen number of replications S can be rather large.

ii) **Estimation under full information**

The same type of approach can be followed when the information set I_t^1 contains asset price changes, and the location and scale functions are parametrized: $\mu(\underline{\Delta p_{t-1}}; \theta), \Omega(\underline{\Delta p_{t-1}}; \theta)$, respectively. The parameter θ is estimated by quasi (pseudo) maximum likelihood:

$$\hat{\theta}_T = \underset{\theta}{Argmax} -\frac{1}{2}\log\det\Omega(\underline{\Delta p_{t-1}}; \theta) - \frac{1}{2}[\Delta p_t - \mu(\underline{\Delta p_{t-1}}; \theta)]'\Omega(\underline{\Delta p_{t-1}}; \theta)^{-1}[\Delta p_t - \mu(\underline{\Delta p_{t-1}}; \theta)].$$

Then the residuals are:

$$\hat{u}_t = \Omega(\underline{\Delta p_{t-1}}; \hat{\theta}_T)^{-1/2}[\Delta p_t - \mu(\underline{\Delta p_{t-1}}; \hat{\theta}_T)],$$

and the multivariate distribution of u_t is approximated by the sample distribution of residuals.

---

[8]We call this simulation approximate because the true parameter θ is replaced by $\hat{\theta}_T$ and the true distribution of errors by the sample distribution of residuals.

We see that in the multivariate framework, the VaR has to be approximated by simulation even at horizon 1. In general, the conditional distribution of $a'\Delta p_t$ given $\underline{\Delta p_{t-1}}$ does not admit a simple analytical form, even if the multivariate distribution of u_t is as simple as multivariate Student, for example.

4.2 Nonparametric Methods

In the nonlinear heteroscedastic autoregressive model, the impact of lagged price changes on risk is captured by the scale function $\sigma(\underline{y_t})$; in particular for small and for extreme values of the error this impact is very close. However, a less constrained specification would be preferred if it allowed to distinguish between the impacts of small and extreme error values on risk. A natural idea seems to be to leave the conditional distribution of price changes (or changes in portfolio values) completely unspecified and proceed with a nonparametric method. Due to the curse of dimensionality, the nonparametric approach can be applied to the changes in portfolio value when they have short memory, that is when the number of autoregressive lags in the polynomial is small. For this reason, we assume in this section that the process (y_t) is stationary Markov of order 1, and we focus on the estimation of the conditional distribution of y_t given y_{t-1}, or equivalently of the joint distribution of (y_{t-1}, y_t).

It is known that the joint cdf of (y_{t-1}, y_t) can be decomposed into the marginal distribution and a term that represents serial dependence. More precisely, we have [Sklar (1959)]:

$$
\begin{aligned}
F_2(y_t, y_{t-1}) &= P[Y_t < y_t, Y_{t-1} < y_{t-1}] \\
&= P[F(Y_t) < F(y_t), F(Y_{t-1}) < F(y_{t-1})],
\end{aligned}
$$

where F denotes the marginal cdf of y_t. Since $F(Y_t)$ follows a uniform distribution on [0,1], we find that:

$$
F_2(y_t, y_{t-1}) = C[F(y_t), F(Y_{t-1})], \tag{4.7}
$$

where C is the joint cdf of $U_t = F(Y_t), U_{t-1} = F(Y_{t-1})$. The function C is called a **copula** cumulative distribution function. Due to the constraint imposed on the marginal distributions of U_t, U_{t-1}, the copula satisfies:

$$
\begin{aligned}
C[u_t, 1] &= u_t, \ \forall u_t, \\
C[1, u_{t-1}] &= u_{t-1}, \ \forall u_{t-1}.
\end{aligned}
$$

The VaR at horizon 1 can be easily expressed in terms of the marginal distribution F and the copula C. Indeed, we have:

$$P[Y_t < VaR_t | Y_{t-1} = y_{t-1}]$$

$$= P[F(Y_t) < F(-VaR_t) | F(Y_{t-1}) = F(y_{t-1})]$$

$$= P[U_t < F(-VaR_t) | U_{t-1} = F(y_{t-1})]$$

$$= \frac{\partial C}{\partial u_{t-1}}[F(-VaR_t), F(y_{t-1})].$$

Thus the VaR is the solution of:

$$\frac{\partial C}{\partial u_{t-1}}[F(-VaR_t), F(y_{t-1})] = \alpha. \qquad (4.8)$$

It seems natural to estimate nonparametrically the functions F and C by their (kernel smoothed) empirical counterparts and then to solve equation (4.8) after replacing the functions F and C by these counterparts. However the difficulty encountered in estimating the empirical quantile described in section (3.1) gets worse in the bidimensional framework and results in inaccurate VaR estimates. Indeed, the rate of convergence of this estimator depends on the dimension of distribution. To circumvent this problem it has been proposed to constrain the copula nonparametrically. For instance, we can consider an **Archimedean copula**. An Archimedean copula is defined by:

$$C(u, v) = \Psi^{-1}[\Psi(u) + \Psi(v)], \qquad (4.9)$$

where Ψ is a real function [9]. Thus, for an Archimedean Copula, serial dependence is captured by the one dimensional function Ψ (instead of the bidimensional function C in the unconstrained case and of the scalar autoregressive parameter ρ in the gaussian case). It is easy to check that:

$$P[C(U_t, U_{t-1}) < s] = s - \Psi(s)/\frac{d\Psi}{ds}(s), \ \forall s. \qquad (4.10)$$

This equality can be used to obtain a consistent functional estimator of the function Ψ. The estimation method consists of three steps.

i) **First step**: We use the data on $y_t, t = 1, ..., T$ to find approximations of the uniformly distributed variables $U_t = F(Y_t)$ in the following way. First, the data are ranked in ascending order $y_{(1)} < ... < y_{(T)}$. Then we assign to each $y_t, t = 1, ..., T$ its rank divided by T. The rank divided by T is called $\hat{u}_t$. It is a value of $\hat{F}(y_t)$, where $\hat{F}$ is the empirical cdf inferred from the data.

---

[9] Ψ has to be the Laplace transform (moment generating function) of a positive random variable [see Joe (1997)].

A similar approach is applied to the lagged values $y_t, t = 0, ..., T - 1$ to derive approximations $\tilde{u}_{t-1}$ for u_{t-1}.

ii)**Second step**: The copula cumulative function evaluated at $\hat{u}_t, \tilde{u}_{t-1}$ can now be approximated by its empirical counterpart:

$$\hat{C}(\hat{u}_t, \tilde{u}_{t-1}) = \frac{1}{T} \sum_{\tau=1}^{T} 1_{\hat{u}_\tau < \hat{u}_t, \tilde{u}_{\tau-1} < \tilde{u}_{t-1}}.$$

iii) **Third step**: By applying formula (4.23), we find a smoothed estimator of the function $A(s) = \Psi(s)/\frac{\partial \Psi}{ds}(s)$ by:

$$\hat{A}(s) = \frac{1}{T} \sum_{t=1}^{T} \Phi \left[\frac{\hat{C}(\hat{u}_t, \tilde{u}_{t-1}) - s}{h} \right],$$

where Φ is the cdf of the standard normal used for smoothing and h is the bandwidth. Then the estimator of function Ψ is derived by integration:

$$\hat{\Psi}(u) = \exp \left[\int^u \frac{1}{s - \hat{A}(s)} ds \right].$$

4.3   Miscellaneous

In this section we review three types of other methods for VaR computation that exist in the literature. They can be based on dynamic specifications other than those considered so far. Also, they may arise as generalizations of some of the approaches described in Section 3.

i) **Switching Regimes**

The idea is to extend the basic gaussian model by allowing for endogeneous switching regimes. Conditional on a given regime the distribution of price changes is multivariate normal. However, when the endogeneous regimes are integrated out,it becomes a mixture of gaussian distributions. This approach accommodates heavy tails, persistence and nonlinear dynamics. More precisely, let us denote by $k = 1, ..., K$ the admissible regimes and by Z_t with values in $\{1, ..., K\}$ the market regime at date t. It is assumed that:

a) (Z_t) is a Markov chain with transition matrix Q.

b) The distribution of price changes Δp_t conditional on $Z_t = k, \underline{\Delta p_t}, \underline{Z_t}$ is multivariate normal $N[\mu_k, \Omega_k]$.

Then the conditional distribution of price changes is:

$$l(\Delta p_t | \underline{\Delta p_{t-1}}) = \sum_{k=1}^{K} p_k(\underline{\Delta p_{t-1}}) N(\mu_k, \Omega_k), \tag{4.11}$$

where $p_k(\underline{\Delta p_{t-1}}) = P[Z_t = k | \underline{\Delta p_{t-1}}]$.

The probabilities p_k can be computed numerically and the parameters μ_k, Ω_k, k varying, and Q can be estimated by means of the Kitagawa's algorithm [see e.g. Hamilton (1989)]. Then the conditional VaR is estimated from drawings in the mixture distribution (4.11), after replacing p_k, μ_k, Ω_k by their estimates [see Billio, Pelizzon (2000) for an application]. Note that this approach is different from the mixture of normal distributions proposed by J.P. Morgan as a new methodology of measuring VaR [Longerstay (1996)]. Under the J.P. Morgan approach, the regime indicators (Z_t) are assumed time independent.

ii) Conditional Autoregressive Value at Risk (CAViaR)

The approach developed by Engle and Manganelli (2001) is a one-dimensional approach, that has to be implemented for each portfolio separately. The basic idea is to write directly a dynamic specification for the VaR, such as:

$$
\begin{aligned}
VaR_t &= \beta_0 + \beta_1 VaR_{t-1} + \beta_2 |y_{t-1}| \\
&= \gamma_0(\beta) + \sum_{j=1}^{\infty} \gamma_j(\beta)|y_{t-j}| \,, \text{say,}
\end{aligned}
$$

where $\gamma_0(\beta), \gamma_j(\beta), j$ varying, are functions of $\beta_0, \beta_1, \beta_2$. The parameter β is estimated by regression quantile [Koenker, Basset (1978)] that is by:

$$
\hat{\beta} = Argmin_{\beta_0,\beta_1,\beta_2} \sum_{t=1}^{T} \{\alpha[y_t - \gamma_0(\beta) - \sum_j \gamma_j(\beta)|y_{t-j}|]^+ + (1-\alpha)[y_t - \gamma_0(\beta) - \sum_j \gamma_j(\beta)|y_{t-j}|]^-\}.
$$

The VaR estimator is:

$$
\widehat{VaR} = \gamma_0(\hat{\beta}) + \sum_{j=1}^{\infty} \gamma_j(\hat{\beta})|y_{t-j}|.
$$

This approach is quite easy to implement, but its drawbacks are twofold. First, the CAViaR models written for different portfolio allocations can be incompatible [see remark in section 4.1]. Second the CAViaR models has to be specified separately for each different risk levels α. This leads to VaR estimates:

$$
\widehat{VaR}(\alpha_0) = \gamma_0(\hat{\beta}_{\alpha_0}) + \sum_{j=1}^{\infty} \gamma_j(\hat{\beta}_{\alpha_0})|y_{t-j}|,
$$

$$
\widehat{VaR}(\alpha_1) = \gamma_0(\hat{\beta}_{\alpha_1}) + \sum_{j=1}^{\infty} \gamma_j(\hat{\beta}_{\alpha_1})|y_{t-j}|,
$$

that do not necessarily satisfy the monotonicity property:

$$\widehat{VaR}(\alpha_0) > \widehat{VaR}(\alpha_1), \text{ if } \alpha_o < \alpha_1.$$

iii) **Local maximum likelihood**

Let us assume that the process (y_t) of changes in portfolio values is a Markov process of order one. Gourieroux, Jasiak (2000) approximate the tail of the true conditional density $l(y_{t+1}|y_t = y)$, say, by a parametric distribution such as a gaussian distribution. The procedure is as follows:

First step: Compute the 1%-empirical quantile from the sample $y_1, ..., y_T$. It is denoted by $\hat{q}$.

Second step: Compute the mean and variance in a neighborhood of $y_{t+1} = \hat{q}$ and $y_t = y$. For the kernel K and bandwidth h, the approximate mean and variance are:

$$\hat{m}(\hat{q}, y) = \sum_{\tau=1}^{T} K\left(\frac{y_\tau - \hat{q}}{h}\right) K\left(\frac{y_{\tau-1} - y}{h}\right) y_\tau / \sum_{\tau=1}^{T} K\left(\frac{y_\tau - \hat{q}}{h}\right) K\left(\frac{y_{\tau-1} - y}{h}\right),$$

$$\hat{\sigma}^2(\hat{q}, y) = \sum_{\tau=1}^{T} K\left(\frac{y_\tau - \hat{q}}{h}\right) K\left(\frac{y_{\tau-1} - y}{h}\right) y_\tau^2 / \sum_{\tau=1}^{T} K\left(\frac{y_\tau - \hat{q}}{h}\right) K\left(\frac{y_{\tau-1} - y}{h}\right) - \hat{m}^2(\hat{q}, y).$$

Third step: Apply the gaussian VaR formula with inputs $\hat{m}(\hat{q}, y_T), \hat{\sigma}^2(\hat{q}, y_T)$ to get:

$$\widehat{VaR}_T = -\hat{m}(\hat{q}, y_T) + \Phi^{-1}(1 - \alpha)\hat{\sigma}(\hat{q}, y_T).$$

5 Porfolio with Derivatives

In financial theory a considerable attention is given to derivative pricing and hedging, especially for derivatives such as European calls written on an underlying asset. Let us recall that a European call with maturity T and strike K will pay $(S_T - K)^+ = Max(S_T - K, 0)$ at date T, where S_T is the price at T of the underlying asset. Since the payoffs of derivatives with any strike, written on the same asset, depend on the same benchmark S_T and are defined by nonlinear payoff functions, people tend to believe that prices of such derivatives are strongly and nonlinearly dependent [10]. Various theories have been developed to support this belief. They provide derivative pricing formulas of two types. Under the complete market hypothesis, the price at t of a European call with strike K and maturity T can be written as:

$$C_t(K, T) = C(S_t, r_t, K, T - t), \tag{5.1}$$

where r_t is the interest rate and C is a deterministic function that depends on the dynamics of the underlying asset price. The Black-Scholes [Black, Scholes (1973)] formula is a typical example of this approach.

---

[10]In reality they also depend on other factors wich influence the demand and supply of derivative assets, especially when these are rather illiquid.

Under incomplete markets, the price can also depend on other factors Z_t that are not the underlying asset prices:

$$C_t(K, T) = C(S_t, Z_t, r_t, K, T - t). \tag{5.2}$$

An example of this approach is the Hull- White model [Hull, White (1987)], in which the unobservable factor is stochastic volatility.

The approaches under both complete and incomplete markets rely on some restrictive assumptions. For example, they assume [11] that a) all assets including the derivatives are liquid, b) they can be traded at any time, c) the derivative prices are functions of state variables (S_t, r_t) or (S_t, Z_t, r_t) only, d) the state variables are Markov processes. However, these assumptions are not satisfied in practice. For example, the index derivatives are written on a market index that is not directly traded on the market [12], the derivative securities are generally not liquid and their prices may depend on some demand and supply effects.

Since derivative trading is a potential cause of financial losses, it is natural to introduce VaR measures for portfolios that include derivatives. Due to the lack of liquidity of such complex assets, it is difficult to come up with a VaR measure based on lagged observed derivative prices, such as for example, the historical simulation method. The challenge of this section is to use the theoretical pricing formulas derived under the liquidity assumption in order to derive reasonably good approximations of the VaR for portfolios with derivatives. The accuracy of such an approximation will depend on the model used by the bank for derivative pricing (called **internal model**, henceforth). Therefore it will be necessary to examine the sensitivity of the VaR with respect to departures from the internal model.

5.1 Parametric Monte-Carlo Method

For ease of exposition let us consider a portfolio of European calls, all written on the same asset. This portfolio is defined by the set of associated strikes and maturities: $(K_i, T_i), i = 1, ..., n$. If one of the strikes is zero, then the portfolio contains the basic asset too. In a complete market framework, the change of the portfolio value is:

$$y_{t+1} = \Delta W_{t+1}(a) = \sum_{i=1}^{n} a_i[C(S_{t+1}, r_{t+1}; K_i, T_i - t - 1) - C(S_t, r_t; K_i, T_i - t)]. \tag{5.3}$$

It is a known function of the current and future interest rates and asset prices. Moreover it depends on the unknown parameter θ that characterizes the dynamics of the price S_t. Let us denote the

---

[11]See e.g. Merton (1974) for a complete list of assumptions for the Black-Scholes model.

[12]Some index mimicking portfolios can actually be traded, such as the SPDR (Standard and Poor Depository Receipts) that mimicks the S&P 500. However the nonlinear dynamic properties of the S&P 500 and of the SPDR are significantly different, especially for extreme values.

function in (5.3) by:

$$y_{t+1} = w_t(S_t, S_{t+1}, r_t, r_{t+1}; \theta). \tag{5.4}$$

The conditional distribution of y_{t+1} has generally no closed-form expression and has to be approximated by Monte-Carlo experiments. Under the assumption of a deterministic interest rate, the procedure is implemented as follows [13]:

i) **First step**: Estimation of the parameter θ

The dynamics of (S_t) is described by the conditional historical distribution of S_t given S_{t-1}: $l(S_t|S_{t-1}; \theta)$, (say). The parameter θ can be estimated from the historical data on (S_t) by the maximum likelihood, for example (the so-called historical approach). The estimate is denoted $\hat{\theta}_T$.

ii) **Second step**: Simulation of future values of S:

For a given value of S_t we can draw simulated values $S_{t+1}^s, s = 1, ..., S$ in the conditional distribution $l(S_{t+1}|S_t; \hat{\theta}_T)$.

iii) **Third step**: Simulation of y_{t+1}

We simulate the values [14] of y_{t+1} as:

$$y_{t+1}^s = w_t(S_t, S_{t+1}^s, r_t, r_{t+1}; \hat{\theta}_T), \ s = 1, ..., S.$$

iv) **Fourth step**: Estimation of the VaR

Finally, the conditional VaR estimate can be derived directly from the empirical quantile of the distribution of $y_{t+1}^s, s = 1, ..., S$.

It is interesting to discuss this approach in the framework of the Black-Scholes model, in which the asset price follows a geometric Brownian motion:

$$dS_t = \mu S_t dt + \sigma S_t dW_t,$$

where (W_t) is a standard Brownian motion. The asset price dynamics depends on two parameters $\theta = (\mu, \sigma)$, whereas the option price depends on the volatility σ only. However, both parameters μ and σ have to be estimated in order to apply the procedure of VaR estimation. Indeed, while the derivative price depends on σ only, its conditional distribution depends on both the volatility σ and the drift μ. This explains why it is necessary to use the historical data $S_t, t = 1, ..., T$ to recover μ in the estimation procedure, rather than use only the data on derivative prices (cross sectional

---

[13]If the interest rate is stochastic, it is also necessary to estimate the dynamics of the rate and to simulate the future interest rates.

[14]In the formula below we assume analytical expressions of the derivative prices. Otherwise they have also to be approximated by Monte-Carlo [see Gourieroux, Jasiak (2001)a, Chapter 11].

or implied volatility approach), since the derivative prices allow for estimation of the volatility σ only.

Finally the approach outlined above can be extended to an incomplete market framework by considering for estimation and simulation the distribution of all state variables, including the unobservable factor Z.

5.2 Taylor Expansion of Nonlinear Portfolios

Approximated closed form expressions of the VaR have also been proposed by the industry to avoid simulation.

i) **The Delta method**

Let us consider the complete market framework. It is possible to build an instantaneously riskless portfolio that consists of a position in a European call and a position in the underlying asset. This riskless position can be reached with an allocation of -1 in the European call and $\delta_t(K, T) = \frac{\partial C}{\partial S}(S_t, r_t, K, T - t)$ in the underlying asset. δ_t is called the **delta** of the derivative security. In this framework, the European call is equivalent to a portfolio in the underlying and riskfree assets with allocations $\delta_t(K, T)$ and $\alpha_t(K, T)$, say, respectively.

Several authors have proposed to apply this result in the following way. Let us consider the initial portfolio of European calls. This portfolio is equivalent to a portfolio including:

$\sum_{i=1}^{n} a_i \delta_t(K_i, T_i)$ units of the underlying asset and a quantity $\sum_{i=1}^{n} a_i \alpha_t(K_i, T_i)$ of the riskless asset. Then the conditional VaR is computed as in section 4 for a linear portfolio in S. It is important to remark that this attractive and simple approach differs from the initial idea of VaR developed by the Basle Committee. According to this idea, the VaR has to be computed for a portfolio with fixed allocations and doesn't take into account portfolio updating during the holding period. In contrast, the δ-method assumes continuous updating of the allocation performed in the optimal way if the internal model is well specified. As a consequence, the δ-based VaR is less than the VaR with constant allocations and underestimates the true VaR. In an extreme case, the internal model views a portfolio with -1 in the derivative and δ_t in the underlying asset as riskfree. Thus, a slight misspecification of the internal model suffices to perceive as riskfree an extremely risky portfolio.

ii) **The delta-gamma method**.

It has been proposed to extend the previous approach by considering a second order Taylor expansion of the derivative price with respect to the price of underlying asset:

$$\Delta C(S_t, r_t, K, T - t) \approx \alpha_t + \delta_t \Delta S_{t+1} + \frac{1}{2} \gamma_t (\Delta S_{t+1})^2, \tag{5.5}$$

where the second order derivative γ_t is the so-called **gamma** of the option. This expansion now includes a nonlinear quadratic function of ΔS. Several authors [see e.g. Jorion (1997) p. 144] proposed to apply the method of the variance-covariance matrix [see 2.2 i)] under the conditional normality of ΔS_{t+1}. The derivative portfolio is such that:

$$\Delta W_t(a) \approx \sum_{i=1}^n a_i \alpha_t^i + (\sum_{i=1}^n a_i \delta_t^i) \Delta S_{t+1} + \frac{1}{2} (\sum_{i=1}^n a_i \gamma_t^i) \Delta S_{t+1}^2. \tag{5.6}$$

We get:

$$
\begin{aligned}
E_t[\Delta W_{t+1}(a)] &\approx \sum_{i=1}^n a_i \alpha_t^i + (\sum_{i=1}^n a_i \delta_t^i) E_t(\Delta S_{t+1}) + \frac{1}{2}(\sum_{i=1}^n a_i \gamma_t^i)[V_t(\Delta S_{t+1}) + (E_t \Delta S_{t+1})^2]; \\
V_t[\Delta W_{t+1}(a)] &\approx (\sum_{i=1}^n a_i \delta_t^i)^2 V_t(\Delta S_{t+1}) + (\frac{1}{2}\sum_{i=1}^n a_i \gamma_t^i)^2 V_t[(\Delta S_{t+1})^2] \\
&\quad + (\sum_{i=1}^n a_i \delta_t^i)(\sum_{i=1}^n a_i \gamma_t^i) Cov_t(\Delta S_{t+1}, \Delta S_{t+1}^2) \\
&\approx (\sum_{i=1}^n a_i \delta_t^i)^2 V_t(\Delta S_{t+1}) + \frac{1}{2}(\sum_{i=1}^n a_i \gamma_t^i)^2 [V_t(\Delta S_{t+1})]^2,
\end{aligned}
$$

since $Cov[\Delta S, (\Delta S)^2] = 0$ and $V[(\Delta S)^2] = 2[V(\Delta S)]^2$ for a gaussian variable. Therefore the first and second order conditional moments of $\Delta W_{t+1}(a)$ are easily computed from the first and second order conditional moments of ΔS_{t+1}. However this approach often used in practice has several drawbacks. First, the expansions are valid when the derivative price is differentiable with respect to S and don't apply to situations close to the expiry date of the option. Second, the mean-variance approach assumes implicitly the approximate normality of the change in portfolio value. Even if ΔS_{t+1} is conditionally normal, this property is no longer satisfied by the portfolio value due to the presence of a quadratic term. Finally the second order Taylor expansion is not properly derived as explained below.

iii) **Linearization of nonlinear portfolios**

Note that the characteristics of a European derivative change in time; in particular the residual maturity decreases while the interest rate varies in time. A proper first order expansion of the derivative price is:

$$
\begin{aligned}
& C(S_{t+1}, r_{t+1}, K, T - t - 1) - C(S_t, r_t, K, T - t) \\
= \; & \frac{\partial C}{\partial S}(S_t, r_t, K, T - t)\Delta S_{t+1} + \frac{\partial C}{\partial r}(S_t, r_t, K, T - t)\Delta r_{t+1} - \frac{\partial C}{\partial T}(S_t, r_t, K, T - t).
\end{aligned}
$$

It differs from the previous first order expansion by the presence of the first order derivatives with respect to the interest rate and to the residual maturity. The same remark holds for a second order Taylor expansion.

The last formula can be extended to include time varying parameters. For example, it is common to use the Black-Scholes model with a time varying volatility σ_t, say. In this case, the expansion is also written with respect to the volatility (or the log-volatility) and involves the associated derivative of the price, that is the **vega** of the option. Then the expansion is easy to use under the assumption of the conditional distribution of $\Delta S_{t+1}, \Delta r_{t+1}, \Delta \log \sigma_{t+1}$ be jointly normal. This approach is suggested by the RiskMetrics Group [see e.g. Malz (2000)], who report variance-covariance matrices, as well as the returns on implied volatilities. However the normality assumption is very unrealistic and the observed implied volatility returns are generally highly leptokurtic and skewed.

iv) **The normality assumption in the case of option prices**

The idea of using the first order expansion is to extend the normality assumption on the price change of the underlying asset to the change in derivative prices. The argument is that the normality is satisfied by "sufficiently large portfolios of independent options", to which the Central Limit Theorem can be applied [Finger (1997)]. However this argument is not valid, since the derivative prices are highly correlated and because the normal approximation is poor in the tails. In fact the conditional distributions of derivative prices are generally far from gaussian. They can admit several modes, feature skewness and fat tails [see e.g. Gourieroux, Jasiak (2001)a, chapter 12].

Despite the aforementioned limitations, the use of delta or delta-gamma methods is recommended as a standard approach by the Capital Adequacy Directive (1993) and by the Banking Supervision Proposal (1995) of the Basle Committee. A survey of the Group of Thirty (1994) showed that 98% out of 125 operators who responded, were using delta or delta-gamma methods. It is important to note that the parametric Monte-Carlo methods don't have the aforementioned drawbacks and nevertheless are easy to implement.

6 Credit Risk

As mentioned in the introduction, the main causes of losses incurred by banks are corporate loans and mortgages. The risk is essentially due to default of payment [15], and thus to the evolution of the solvency of a borrower. Two features have to be taken into account in a study of credit portfolio. These are the heterogeneity of borrowers and contracts, and the lack of liquidity for the majority of loans. Indeed, only a fraction of loans can be traded on secondary markets and have market prices; these are generally corporate bonds and mortgage backed securities. The other types of risky credits are mortgages, consumption loans, revolving credit (credit cards), over-the-counter

---

[15]and also to prepayments of mortgages. The prepayment risk is not discussed in this section.

(OTC) corporate loans, cash advances, known as **retail credits**.

In the first section, we discuss the link between the distribution of default and the actuarial value (resp. market price) of an OTC loan (resp. bond). In the second section we discuss the assessment of past default rates from i) the data on individual behavior of borrowers, ii) observations on bonds, iii) observations on equity prices. The credit migration approach is described in the third section. Finally, the last section contains the analysis of the profit and loss distributions for portfolios of either bonds, or retail loans.

6.1 Spread of Interest Rates

To set the stage for a discussion on the relation between prices and default probability, we first consider a consumer loan with an initial balance B_0 and a fixed interest rate r, to be repaid in H units of time by means of constant monthly payments of amount m. If the borrower has a zero probability of default, the following actuarial relationship holds:

$$B_0 = \frac{m}{1+r} + \frac{m}{(1+r)^2} + \cdots + \frac{m}{(1+r)^H},$$ (6.1)

which equates the initial balance to the sum of discounted cash-flows. This relation can be used to find the balance B_0 that corresponds to given m, r, H, or to find the rate r for given B_0, m, H.

The actuarial formula (6.1) needs to be modified when the probability of default is different from zero. Let us denote by Y the time to default (with the time origin $Y = 0$ set at the date when the credit is granted), and assume that after Y the borrower will not repay, even a fraction of the remaining balance (i.e. the recovery rate is equal to zero) [16]. Then, the actuarial convention implies:

$$
\begin{aligned}
B_0^* &= \frac{m^*}{1+r^*}P[Y \geq 1] + \cdots + \frac{m^*}{(1+r^*)^H}P[Y \geq H] \\
&= m^* \left[\frac{S(1)}{1+r^*} + \cdots + \frac{S(H)}{(1+r^*)^H} \right],
\end{aligned}
$$ (6.2)

where S denotes the survivor function for time to default. Formulas (6.1), (6.2) can be compared in two different ways.

i) If the rate $r^* = r$ and the monthly payment $m^* = m$ are given, we get different actuarial values for the loan depending on the presence of potential default. The value computed without default risk, that is $B_0 = \sum_{h=1}^{H} \frac{m}{(1+r)^h}$, is strictly larger than the value $B_0^* = \sum_{h=1}^{H} \frac{mS(h)}{(1+r)^h}$. The omission of default risk causes overevaluation of the credit portfolio.

---

[16] The magnitude and timing of the recovery should also be taken into account. For ease of exposition, we assume a zero recovery rate. Even though this assumption is unrealistic, it is important to note that it is used by the markets to recover the implied probability of default from bond prices. Moreover, when the recovery is assumed independent of default, the actuarial prices are simply inferred from the proposed ones by multiplying by the expected recovery rate. This approach is recommended by the Basle Committee.

ii) If the value $B_0^* = B_0$ and the monthly payment $m^* = m$ are given, we get different rates that satisfy the actuarial conditions. It is easy to check that the rate r is strictly higher than the rate r^* to compensate for default risk. The difference $s = r - r^*$ is called the **spread** of interest rate.

As an illustration , let us assume an exponentially distributed time to default Y with default intensity λ. We get $S(h) = \exp(-\lambda h)$ and,

$$
\begin{aligned}
B_0 &= m\left[\frac{S(1)}{1+r^*} + \cdots + \frac{S(H)}{(1+r^*)^H}\right] \\
&= m\left[\frac{\exp -\lambda}{1+r^*} + \cdots + \frac{\exp -\lambda H}{(1+r^*)^H}\right] \\
&= m\left[\frac{\exp -\lambda}{1+r^*} + \cdots + \left[\frac{\exp -\lambda}{1+r^*}\right]^H\right].
\end{aligned}
$$

We find that

$$
1 + r = (1 + r^*)\exp \lambda
$$

$$
\Leftrightarrow \quad s = r - r^* = (1 + r^*)[\exp \lambda - 1].
$$

The spread is an increasing function of default intensity λ.

Until now, the approach assumed a constant rate of interest and a flat term structure. It can easily be extended to any type of fixed income bonds (or retail loans without indexed payments) and to a varying term structure of interest rates. Let us consider a bond with known future payments F_τ (say) at dates τ, and denote by $B(t, t + h)$ the price at t of the zero coupon bond that pays 1 \$ at date $t + h$. Without default risk, the price of this bond at date t is:

$$
P_t(F) = \sum_{h=1}^{\infty} F_{t+h} B(t, t + h). \tag{6.3}
$$

The price formula is derived by observing that a fixed income bond is a portfolio of zero coupon bonds and by applying the arbitrage free condition. In the presence of default risk due to the borrower, the price of bond will decrease. If default is independent of the evolution of the riskfree interest rate, the price of bond with default for a risk neutral investor is:

$$
P_t(F, S) = \sum_{h=1}^{\infty} F_{t+h} B(t, t + h) S_t(t, t + h), \tag{6.4}
$$

where $S_t(t, t + h) = P_t[Y \geq t + h | Y \geq t]$ and the time to default is measured since the time origin. The conditioning is necessary because the bond can only be priced for a contract while it is still alive, and the index t means that the information set used to predict Y increases.

Formulas (6.3) and (6.4) involve two term structures of interest rates: the term structure without default risk is characterized by the set $B(t, t + h)$, $h = 1, ..., H$; the term structure with default risk is characterized by $B^*(t, t + h) = B(t, t + h)S_t(t, t + h)$, $h = 1, ..., H$, and depends on the distribution of time to default. $S_t(t, t + h)$, $h = 1, ..., H$, defines the term structure of spread that is the mapping $h \rightarrow s_{t,t+h} = r_{t,t+h} - r_{t,t+1}^* = \frac{1}{h} \log \frac{B(t,t+h)}{B^*(t,t+h)} = -\frac{1}{h} \log S_t(t, t + h)$ [17].

6.2   Assessment of Default Rates

There are two sources of randomness in the future price of a bond or a retail loan. Therefore we have to predict the future riskfree term structure and the future probability of default. Since these features are generally treated independently we will focus on the estimation of past probabilities $S_t(t, t + h)$. There exist two approaches, that differ with respect to the assumption on the loans being traded on a secondary market.

Recall that typically the bond market trades corporate bonds offered by some thousands of issuers. The issuings are regular, most of them standardized and provided with information on the rating of the firm and its balance sheet. On the other hand, retail loans are, for example, retail consumer loans or mortgages for several millions of borrowers held in the portfolio of a given bank. In general, the contracts concern small amounts and are very heterogeneous with respect to the initial balance,maturity, interest rate, pattern of monthly payments and characteristics of the borrower.

   i)**Recovering default rates from market bond prices**

Let us consider a given corporation j that issued bonds $l = 1, ..., L_{jt}$, which are traded on the market at date t. At date t these bonds differ with respect to their payoff patterns F_{t+h}^{jl} and prices $P_t^{jl}(F, S)$. They are related by:

$$P_t^{jl}(F, S) = \sum_{h=1}^{\infty} F_{t+h}^{jl} B(t, t + h) S_t^j(t, t + h), \ l = 1, ..., L_{jt}, \tag{6.5}$$

where the conditional survivor function depends only on the borrower j and not on the bond.

Let us assume that the number L_{jt} of traded bonds is large and that their cashflow patterns are very diversified. Then we can apply a standard approach to recover the term structure of these corporate bonds, such as the regression method, local polynomials or splines, which belongs to the smoothing techniques used for recovering the term structure of Treasury bonds. This procedure yields the approximated term structure of firm j:

$$\hat{B}^{*j}(t, t + h) \approx B(t, t + h) S_t^j(t, t + h), \text{ for any } t, h.$$

---

[17]Similarly, when the recovery rates are taken into account, there exists a term structure of recovery rates.

Then by using an estimated riskfree term structure $\hat{B}(t, t+h), t, h$ varying, based on the T-bond prices, we find the estimators of individual default probabilities [18]:

$$\hat{S}_t^{*j}(t, t+h) = \hat{B}^{*j}(t, t+h)/\hat{B}(t, t+h). \tag{6.6}$$

This approach requires only the knowledge of bond prices recorded on the market. Therefore it can be used by banks that don't own the data on individual credit histories of customers. Practice shows that the market based information is rather poor and leads to biased estimators of past rates of default. This is easy to explain since, for a given issuer j, the number of corporate bonds that are actively traded on the market at a given date is limited. For this reason the quoted prices can be quite different from the theoretical values and the standard approach outlined above fails. To partly circumvent this difficulty firms can be classified in homogenous categories $k = 1, .., K$. Typically, the categories are defined according to the actual Standard and Poor (resp. Moody's) rating of each firm, starting from the highest rating AAA (resp. Aaa), to the lowest one CCC (resp Caa) [19]. Then, the term structure of spread is assumed identical for all issuers in the same rating category. This allows to use a larger number of traded bonds to recover the past rates of default. This approach is discussed in greater detail in section 6.3.

ii) **Recovering default rates from equity prices**

It has also been proposed to use the Merton's model [Merton (1974), Crosbie (1998), Janosi, Jarrow, Yildrim (2001)] to recover default probabilities from data on equity value. More precisely, let us consider a given company j with equity value, firm asset value and liabilities denoted by $V_{E,t}, V_{A,t}$ and L_t, respectively. Then the equity can be considered a call option on the future value $V_{A,t+1}$ with strike L_{t+1}. If the liabilities are predetermined and the asset values follow a Black-Scholes model, than the value $V_{E,t}$ can be derived by the Black-Scholes option pricing formula as a function of $V_{A,t}$ and volatility σ_A of the asset value. Moreover, under the Black-Scholes model, the equity and asset volatilities are related by:

$$\sigma_E V_{E,t} = V_{A,t}\sigma_A \delta_t,$$

where δ_t is the delta of the call option. Therefore, given the data on equity value and equity volatility, we can find σ_A and $V_{A,t}$ by using the last equality and the Black-Scholes option pricing formula, respectively. The results allow further computation of the conditional probability of default at $t + h$:

---

[18]called the implied survivor probability.

[19]The complete list of ratings assigns one of the following 10 symbols:

Moody's: Aaa, Aa, A, Baa, Ba, B, Caa, Ca, C, D.

Standard and Poor: AAA, AA, A, BBB, BB, B, CCC, CC, C, D.

$$S_t(t, t+h) = P[V_{A,t+h} < L_{t+h}, V_{A,t+h-1} > L_{t+h-1}, ..., V_{A,t+1} > L_{t+1}|V_{A,t}].$$

This approach has been recommended by the KMV corporation. It can be criticized for disregarding the information contained in bond prices and for assuming the future liabilities known, which is equivalent to disregarding the possibility of future borrowing and debt renegotiation. As well, the method is very sensitive with respect to the continuous time model of the firm value.

iii) **Recovering default rates from individual credit histories**

Let us now consider the retail loans. To eliminate a large part of incompleteness due to individual (contract) heterogeneity, we first define homogenous categories of the same type of contracts (that is with identical initial balance, term, interest rate, contractual pattern of monthly payments) and similar individuals, with almost the same attitude towards default (and prepayment). These categories are indexed by $k, k = 1, \ldots, K$. Then these categories can be partitioned with respect to the generation of loans, leading to a set of cohorts doubly indexed by k, τ, where k is the category index and τ the starting date of contract. If the number of contracts in each cohort is sufficiently large (greater than 200-300), we can eliminate a large part of incompleteness by averaging out homogenous contracts. In such a case, we essentially consider aggregate data on default rates (prepayment rates, recovery rates) cohort by cohort. To simplify exposition, we focus on default and don't consider eventual prepayments or partial recoveries.

Let us consider a time unit of one semester. For each cohort, we observe default rates over all semesters between the starting date of credit and the current date. We denote by $D_k(\tau; h)$ these rates for cohort k, τ at semester $\tau + h$; h denotes the age of contract, that is the time elapsed since the agreement was signed. For each category k, we get a double entry table, which may contain various pairs of entries, such as the generation and current date, the generation and age, or the current date and age. For illustration, we show in Tables below data on loans with maturity equal to two years. Accordingly the maximal age is 4 semesters. The starting date indicates the year and semester. The index of each category is not given. The most recent semester observed is (the end of) 99.2. In practice, Tables 6.2 and 6.3 are easier to read and require less memory for computer storage [20].

---

[20]Similar tables can be designed for the recovery rates.

Table 6.1 : Default rate by generation and current date

current date generation	97.1	97.2	98.1	98.2	99.1	99.2
97.1	$D(97.1;1)$	$D(97.1;2)$	$D(97.1;3)$	$D(97.1;4)$		
97.2		$D(97.2;1)$	$D(97.2;2)$	$D(97.2;3)$	$D(97.2;4)$	
98.1			$D(98.1;1)$	$D(98.1;2)$	$D(98.1;3)$	$D(98.1;4)$
98.2				$D(98.2;1)$	$D(98.2;2)$	$D(98.2;3)$
99.1					$D(99.1;1)$	$D(97.1;2)$
99.2						$D(99.2;1)$

Table 6.2 : Default rate by generation and age

age generation	1	2	3	4
97.1	$D(97.1;1)$	$D(97.1;2)$	$D(97.1;3)$	$D(97.1;4)$
97.2	$D(97.2;1)$	$D(97.2;2)$	$D(97.2;3)$	$D(97.2;4)$
98.1	$D(98.1;1)$	$D(98.1;2)$	$D(98.1;3)$	$D(98.1;4)$
98.2	$D(98.2;1)$	$D(98.2;2)$	$D(98.2;3)$	
99.1	$D(99.1;1)$	$D(99.1;2)$		
99.2	$D(99.2;1)$			

Table 6.3 : Default rate by age and current date

current date age	97.1	97.2	98.1	98.2	99.1	99.2
1	$D(97.1;1)$	$D(97.2;1)$	$D(98.1;1)$	$D(98.2;1)$	$D(99.1;1)$	$D(99.2;1)$
2		$D(97.1;2)$	$D(97.2;2)$	$D(98.1;2)$	$D(98.2;2)$	$D(99.1;2)$
3			$D(97.1;3)$	$D(97.2;3)$	$D(98.1;3)$	$D(98.2;3)$
4				$D(97.1;4)$	$D(97.2;4)$	$D(98.1;4)$

6.3   The Credit Migration Approach

Due to heterogeneity of bond issuers in section 6.2i) we distinguished conditional distributions of time to default for each category of rating. The inconvenience is that the approach needs to be applied sequentially at each date and regularly updated. The objective of the credit migration approach is to analyze the joint dynamics of rating and default.

   i) **The model**

The model was initially conceived as a continuous time model and introduced by Jarrow, Lando and Turnbull (1997). In this section we present an extended and discrete time version of the original model. The key assumption is the existence of a finite number of states $k = 1, ..., K$ that represent risk quality. At each date the borrower occupies one state, and at a future date he can stay or migrate to another state. We denote by (Z_t) the qualitative process formed by the sequence of states occupied by the borrower. In general, it is assumed that this process is a Markov chain with a transition matrix $Q = (q_{kl})$, with elements $q_{kl} = P[Z_t = k | Z_{t-1} = l]$.

The knowledge of recent state history is assumed to be sufficient to define the term structure of credit spread. More precisely, if the borrower has spent h periods of time in state k, after a transition from state l into k, the credit spread is captured by the value of survivor function $S_{kl}(h)$ indexed by k, l. Thus the model is parametrized by the transition matrix Q and the set of survivor functions S_{kl}, k, l varying. To illustrate the zero coupon price dynamics, let us consider a given risk state history. The term structures with and without default risk are given in Table 6.4:

Table 6.4: The Term Structures

date t	Z_t	$B(0,t)$	$B^*(0,t)$
1	1	$B(0,1)$	.
2	3	$B(0,2)$	$B(0,2)S_{31}(1)$
3	3	$B(0,3)$	$B(0,3)S_{31}(2)$
4	3	$B(0,4)$	$B(0,4)S_{31}(3)$
5	1	$B(0,5)$	$B(0,5)S_{13}(1)$
6	2	$B(0,6)$	$B(0,6)S_{21}(1)$
7	2	$B(0,7)$	$B(0,7)S_{21}(2)$
8	3	$B(0,8)$	$B(0,8)S_{32}(1)$
9	3	$B(0,9)$	$B(0,9)S_{32}(2)$
10	3	$B(0,10)$	$B(0,10)S_{32}(3)$

Some versions of the model appeared in financial literature under simplified assumptions. For example, Jarrow, Lando, Turnbull (1997) assumed that the spread is a constant function of the current state only: $S_{kl}(h) = \exp(-\lambda_k h)$. Longstaff, Schwartz (1995), Duffie, Kan (1998), Lando (1998) allowed for more complicated term structure patterns, but assumed also dependence on the current state only. Intuitively, it is clear that the past state contains information about future default too. Loosely speaking, we don't expect to observe the same spread term structure for a borrower with AA rating who was AAA before and for another borrower with the same current AA state, but who was rated A before.

ii) **Statistical inference when the state is observable**

Let us assume independent risk dynamics for different borrowers. When the state histories are observed, the transition matrix is easily estimated and replaced by its empirical counterpart. Then we can consider all the observed histories with a spell in k after a transition from l, and infer an estimator of the survivor function S_{kl} from the observed bond prices that pertain to these spells.

This approach is followed by market practitioners with admissible states defined as ratings AAA, AA, A, BBB, BB, B, CCC determined by the Standard and Poor [21]. For this purpose, the consulting firms report regularly the estimated transition matrix and spread term structures at horizon of one year that depend only on the current state k [see Tables 6.5 and 6.6].

Table 6.5: Estimated Transition Matrix

rating	AAA	AA	A	BBB	B	BB	CCC	Default
AAA	90.81	8.33	0.68	0.06	0.12	0	0	0
AA	0.70	90.65	7.79	0.64	0.06	0.14	0.02	0
A	0.09	2.27	91.05	5.52	0.74	0.26	0.01	0.06
BBB	0.02	0.33	5.95	86.93	5.30	1.17	0.12	0.18
BB	0.03	0.14	0.67	7.73	80.53	8.84	1.00	1.06
B	0	0.11	0.24	0.43	6.48	83.46	4.07	5.20
CCC	0.21	0	0.22	1.30	2.38	11.24	64.86	19.79

---

[21] Note that the Standard and Poor rating and the Mooody's rating are not completely compatible, especially for dates close to a change of rating, that is to a sudden change of risk level [see,e.g. the discussion in Kliger, Sarig (2000)].

Table 6.6: The Spread Curve

Category	Year 1	Year 2	Year 3	Year 4
AAA	3.60	4.17	4.73	5.12
AA	3.65	4.22	4.78	5.17
A	3.72	4.32	4.93	5.32
BBB	4.10	4.67	5.25	5.63
BB	5.55	6.02	6.78	7.27
B	6.05	7.02	8.03	8.52
CCC	15.05	15.02	14.03	13.52

We observe that the spread is not constant. It increases generally with the term, but may decrease for low rating. Indeed, the long term spread takes into account the fact that the contract is still alive which is a very positive information on a priori risky borrowers.

iii) Unobservable States

The previous approach is appealing since it is easy to implement and to understand, and capable to accomodate the joint dynamics of rating and default. However this approach identifies the risk category with a rating assigned by a private company, which has a time varying rating strategy. For example, the transition matrices reported by Moody's are clearly time varying and seems to reflect their tendency to tighten the rules of evaluating borrowers. Moreover, even though the details on how the Moody's ratings are computed are not exactly known for confidentiality reason, we can expect that the ratings depend not only on the structure and dynamics of the balance sheet, but also on the market price history. This dependence is not compatible with the assumption on the evolution of Z_t as a Markov chain, and the dependence of the current state on the last state only.

When the state is considered unobservable, the model becomes a complicated Hidden Markov model, which requires simulation based methods for estimation and recovering the underlying state histories. To our knowledge such an estimation with unobservable states has not been yet performed.

6.4   VaR for Credit Portfolio

There exist various approaches to the analysis of profit and loss distribution and determination of the VaR. We describe below two of them. The first approach is based on implementation of the credit migration approach for prediction making and is suitable for corporate bonds. The second one assumes that heterogeneity is exogenous, and is more appropriate for retail consumer loans. Finally, we discuss the problem of default correlation.

i) The future portfolio value

Before determining the VaR, we need to define precisely the future value of portfolio that contains bonds. Indeed, for a given bond, there is a time varying cash-flow pattern and also

possibly coupon payments during the holding period. To simplify the discussion, let us consider a single bond with a cash-flow sequence $F_\tau, \tau = t+1, t+2, ...$ Without default risk its price at t is:

$$W_t = P_t(F) = \sum_{h=1}^{\infty} F_{t+h} B(t, t+h). \tag{6.7}$$

With default risk, its price becomes:

$$W_t = P_t(F, S) = \sum_{h=1}^{\infty} F_{t+h} B(t, t+h) S_t(t, t+h). \tag{6.8}$$

Its future value at $t+2$, say, will depend on the coupons which have been paid at dates $t+1$ and $t+2$, and on the value at $t+2$ of the residual bond with cash-flows $F_\tau, \tau \geq t+3$. To aggregate these components, we have to explain how the cashed-in coupons are reinvested. To simplify the exposition we assume that they are invested in the riskfree asset.

Without default risk the future value of this bond is:

$$W_{t+2} = F_{t+1}[B(t+1, t+2)]^{-1} + F_{t+2} + \sum_{h=1}^{\infty} F_{t+2+h} B(t+2, t+2+h).$$

With default risk the future value of this bond can be:

$W_{t+2} = 0$, if there is default at $t+1$;

$W_{t+2} = F_{t+1}[B(t+1, t+2)]^{-1}$, if there is default at $t+2$;

$W_{t+2} = F_{t+1}[B(t+1, t+2)]^{-1} + F_{t+2} + \sum_{h=1}^{\infty} F_{t+2+h} B(t+2, t+2+h) S_{t+2}(t+2, t+2+h)$,

otherwise.

At date t this future value is stochastic since i) we don't know if the credit agreement will still be alive at dates $t+1$ and $t+2$. ii) the future term structure is unknown, iii) the conditional survivor probabilities have to be updated.

ii)**The credit migration approach**

The credit migration model is convenient for approximation of the profit and loss distribution by simulation. Before explaining the procedure, let us first describe in detail the credit portfolio. The portfolio contains the bonds of n issuers $i = 1, .., n$. Each issuer i is characterized by his state history $\underline{Z}_{i,t} = (Z_{i,t}, Z_{i,t-1}, ...,)$; the different bonds of the same issuer i included in the portfolio can be aggregated leading to a sequence of aggregated cash-flows $F_{i,\tau}, \tau = t+1, t+2, ...$ We denote by $P_{i,t}(\underline{Z}_{i,t})$ the price at t of this sequence of cash-flows, and by $P_{i,t+h}(\underline{Z}_{i,t+h})$ the price at $t+h$ of the residual sequence $F_{i,\tau}, \tau \geq t+h+1$. These prices include the cost of default risk and depend on the individual state history, which influences the evaluation of default probabilities.

The current value of the credit portfolio is:

$$W_t = \sum_{i=1}^{n} W_{i,t} = \sum_{i=1}^{n} P_{i,t}(\underline{Z}_{i,t}),$$

where $W_{i,t}$ denotes the total value of all bonds of issuer i in the portfolio. The future value is:

$$W_{t+h} = \sum_{i=1}^{n} W_{i,t+h}.$$

The conditional distribution of this future value can be approximated by the sample distribution of:

$$W_{t+h}^s = \sum_{i=1}^{n} W_{i,t+h}^s, \ s = 1, ..., S, \tag{6.9}$$

where the simulated values for each issuer $W_{i,t+h}^s$, i varying, are drawn independently. Let us now describe the drawing of the issuer specific value at horizon 2, say. At date t, after a transition from state l_t, the issuer i stays in state $Z_t = k_t$ for time H_t.

First step: Drawing of the next state Z_{t+1}

The next state Z_{t+1}^s is drawn in the conditional distribution of Z_{t+1} given $Z_t = k_t$ by using the estimated transition matrix.

Second step: Simulation of survival at date $t+1$

Two cases have to be distinguished depending on whether $Z_{t+1}^s = Z_t$.

If $Z_{t+1}^s = Z_t = k_t$, there is default at $t+1$ with probability $1 - S_{k_t, l_t}(H_t + 1)/S_{k_t, l_t}(H_t)$ and no default otherwise.

If $Z_{t+1}^s = k_{t+1} \neq k_t$, there is default at $t+1$ with probability $1 - S_{k_{t+1}, k_t}(1)$ and no default otherwise.

These distributions are used to simulate the potential default at date $t+1$.

Third step: Drawing of the state Z_{t+2}.

This step is applied provided that the contract is still alive. The state is drawn in the conditional distribution of Z_{t+2} given $Z_{t+1}^s = k_{t+1}$, where k_{t+1} is the state drawn at step 1.

Fourth step: Simulation of time to default at $t+2$.

Three cases have to be distinguished. They are described below along with the associated conditional probability of default.

case 1: $Z_{t+2}^s = Z_{t+1}^s = Z_t = k_t$

probability of default: $1 - S_{k_t, l_t}(H_t + 2)/S_{k_t, l_t}(H_t + 1)$;

case 2: $Z_{t+2}^s = Z_{t+1}^s = k_{t+1} \neq Z_t = k_t$

probability of default: $1 - S_{k_{t+1}, k_t}(2)/S_{k_{t+1}, k_t}(1)$;

case 3: $Z_{t+2}^s = k_{t+2} \neq Z_{t+1}^s = k_{t+1}$

probability of default: $1 - S_{k_{t+2}, k_{t+1}}(1)$.

The simulated time to default is denoted by Y^s.

Fifth step: The simulated issuer specific value is computed from:

$$W_{i,t+2}^s = F_{i,t+1}[B(t,t+1,t+2)]^{-1}1_{Y_i^s>t+1} + F_{i,t+2}1_{Y_i^s>t+2} + 1_{Y_i^s>t+2}P_{i,t+2}(\underline{Z}_{i,t+2}^s),$$

where $\underline{Z}_{i,t+2}^s = (Z_{i,t+2}^s, Z_{i,t+1}^s, \underline{Z}_{i,t})$ and $1_{Y_i^s>t+1} = 1$ if $Y_i^s > t+1$, 0 otherwise, denotes the indicator function.

This approach assumes that the future riskfree term structure is known and so is the price $P_{i,t}(\underline{Z}_{i,t})$ when the price history is given. Let us focus on default risk and disregard the risk on T-bond interest rate. The prices $P_{i,t}(\underline{Z}_{i,t})$ [or $P_{i,t+2}(\underline{Z}_{i,t+2}^s)$] are functions of the probabilities of default. These probabilities are unknown and can be approximated by simulations that take into account future risk migration [22]. Let us replace the prices in the last expression by their approximations $\hat{P}$, say. The change in portfolio value becomes:

$$\Delta W_{t+2}^s = \sum_{i=1}^{n} \{F_{i,t+1}[B(t+1,t+2)]^{-1}1_{Y_i^s>t+2} + 1_{Y_i^s>t+2}\hat{P}_{i,t+2}(\underline{Z}_{i,t+2}^s) - \hat{P}_{i,t}(\underline{Z}_{i,t})\}\ s = 1,...,S.$$

$$(6.10)$$

Then the VaR is the empirical α-quantile of the distribution of ΔW_{t+2}^s, $s = 1,...,S$.

Finally note that the estimated transition matrices admit generally coefficients on the main diagonal that are all close to 90 % (see Table 6.5). It is often suggested [see e.g. CreditMetrics] to avoid Monte-Carlo computation of the prices $P_{i,t}(\underline{Z}_{i,t})$ [resp. $P_{i,t+2}(Z_{i,t+2}^s)$] by assuming that after t [resp. after $t+2$] no migration between risk categories will take place. This crude approximation greatly simplifies determination of the VaR, but can induce significant bias. To clarify this, let us consider the most risky category CCC. The computation of the bond price, as if the issuer were to stay in the same rating category CCC, disregards the high probability of default and the possible zero price of the bond in the future.

iii) **The cohort approach**

Let us now consider the retail loans. We introduce a dynamic model for default rates aggregated by cohorts, which is easy to estimate and simulate. Various specifications can be used to predict the future default rates. We just give an example, that includes autoregressive effects of lagged default, macroeconomic factors and allows for unobserved time heterogeneity. Due to the autoregressive part, the specification for the first semester of a loan agreement of any maturity is different from the specification for the next semester and the following ones. For the first semester of the loan, we have no information on past default history of the cohort, but it can be approximated by using the basic score of credit granting used by the credit institution. We denote by $S_{k,\tau}$ and $\sigma_{S,k,\tau}^2$ the average basic score and its dispersion for the cohort k,τ. For semester $h=1$, we use the following

---

[22]It is natural to approximate these prices by simulations because they can be viewed as prices of American options.

logistic model:

$$l[D_k(\tau;1)] = a_1 + b_1 l(S_{k,\tau}) + c_1 \sigma_{S,k,\tau}^2 + d_1' X_{\tau+1} + \alpha_1 l[D_k(\tau-1;1)] + \varepsilon_k(\tau;1), \qquad (6.11)$$

where the components of X are macroeconomic variables, $\varepsilon_k(\tau;1)$ is an error term and $l(x) = \log[x/(1-x)]$ denotes the logit transformation. For next semesters of the loan, we can introduce an additional autoregressive effect associated with the same cohort (τ, k) and a lagged effect of the previous cohort $(\tau-1, k)$:

$$
\begin{aligned}
l[D_k(\tau;h)] &= a_h + b_h l(S_{k,\tau}) + c_h \sigma_{S,k,\tau}^2 + d_h' X_{\tau+h} + \alpha_h l[D_k(\tau-1;h)] \\
&+ \beta_h l[D_k(\tau, h-1)] + \varepsilon_k(\tau, h), h \geq 2. \qquad (6.12)
\end{aligned}
$$

The joint model (6.11), (6.12) is a spatial regression model. It is completed by specifying the distribution of the error terms $\varepsilon_k(\tau, h)$, for any k, τ, h. We assume independence between cohorts and a possibility of correlation between semesters. More precisely we assume :

$[\varepsilon_k(\tau, h), h = 1, \ldots, H]$, τ, k varying, are independent, normally distributed, with zero mean and variance-covariance matrix Σ.

The parameters $a_h, b_h, \ldots, \beta_h, h = 1, \ldots, H$ and Σ can be estimated by ordinary least squares. Even though the number of parameters is large, we have for each semester h a number of observations equal to the number of generations times the number of categories.

The estimated models are used for prediction making, in particular for finding by Monte-Carlo the next columns of Table 6.3. For example, for the future date 00.1 (first calendar semester of year 2000) an error $\epsilon^s(00.1; 1)$ is drawn and the simulated default rate $D^s(00.1; 1)$ is determined by model (6.11) from $D(99.2; 1)$ and the simulated error. For the second row of that column we simulate the error $\epsilon^s(99.2; 2)$ and $D^s(99.2; 2)$ is determined by model (6.12), $D(99.2; 1)$, $D(99.1; 1)$ and $\epsilon^s(99.2; 2)$, and so on. Note the difficulty in predicting the future values of macroeconomic variables X. A solution consists of considering several scenarios of their evolution to assess the default rate.

iv) **Default correlation**

The previous procedures assumed independence of risks of various borrowers and disregarded the possibility of simultaneous bankruptcies. Under the migration approach simultaneous bankruptcies can be examined by considering more complicated transition matrices that represent, for example, the joint migration probabilities of two issuers, rated BB and A. Simultaneous bankruptcies can be incorporated in the cohort approach too by allowing for correlation between the error terms

of two different generations or categories. These model extensions are difficult to implement, especially since they involve complicated multivariate distributions. The challenge consists of finding a constrained multivariate distribution that would provide good fit to the data and be relatively easy for prediction making. Currently, various parametric families of copulas and factor models are topics of ongoing research [see e.g. Schonbucher(2000), Gourieroux, Monfort(2002)].

7   Future Directions for Research and Development

In previous chapters, we have described various approaches developed in the academic and professional literature to determine the Value at Risk. They can be applied to portfolios of liquid financial assets, portfolios of derivatives, and also can take into account the risk of default. However, work in this field is far from completion. The aim of this chapter is to provide some insights on various promising directions for future research.

7.1   Coherent Risk Measures

Despite its successs, the Value at Risk defined as a conditional quantile differs from the risk measures usually employed in the insurance industry. The reason is that it disregards the magnitude of loss when it occurs. To correct for this drawback, Artzner, Delbaen, Heath (1997) proposed, in a two period framework, a constructive approach to compute the capital requirement. They introduced four axioms given below.

Let us denote by $R_t(W)$ the required reserve amount for the future portfolio value W. The axioms concern the properties of monotonicity, invariance with respect to drift, homogeneity and subadditivity.

(i) **Monotonicity**

If W is less preferable than W^* for the stochastic dominance at order 1 (that is if the cumulative distribution function of W^* is larger than the cumulative distribution function of W), then

$$R_t(W) \geq R_t(W^*).$$

(ii) **Invariance with respect to drift**

$$R_t(W + c) = R_t(W) - c,$$

for any W and any deterministic amount c.

(iii) **Homogeneity**

$$R_t(\lambda W) = \lambda R_t(W), \ \forall \lambda \geq 0, \forall W.$$

(iv) **Subadditivity**

$$R_t(W + W^*) \leq R_t(W) + R_t(W^*), \ \forall W, W^*.$$

Homogeneity and subbadditivity imply the convexity of the function R_t. It is easy to check that the conditional quantile does not satisfy the convexity condition. Artzner et alii (1997) have described all functions R_t, the so called **coherent risk measures** that satisfy the four axioms, and given their interpretations in terms of expected utility. In particular, they show that the expected shortfall or **Tail VaR**:

$$TVaR(a, \alpha) = E_t[W_{t+1}(a)|W_{t+1}(a) - W_t(a) + VaR_t(a, \alpha) < 0], \tag{7.1}$$

is a coherent risk measure. TVaR measures the expected value of the portfolio conditional on a loss of probability α. It can be considered as the (historical) price of a reinsurance contract and the capital requirement is viewed as a self-reinsurance.

The axiomatic approach is useful as a basis for discussion about the nature of risk measure. Typically, it emphasizes the importance of the size of loss, and not only of its occurrence. However, it can be criticized for the following reasons.

i) Even if the (conditional) quantile function doesn't satisfy the convexity property for any portfolio value, this property can be satisfied for the conditional distribution of returns and the portfolio allocations, which are encountered in practice [see e.g. Gourieroux, Laurent, Scaillet (2000)].

ii) The homogeneity and subadditivity axioms are clearly not satisfied in practice. Indeed the price per share depends on the traded quantities. For example, it decreases with the quantity for a sell transaction. This effect is especially strong, when the market is close to a crash. This stylized fact is not compatible with axioms (iii) and (iv), which assume that by increasing the size of portfolio risk is diminished.

Moreover, if a coherent risk measure had been selected by the regulators, the banks would have interest in merging to diminish the amount of required capital (due to axioms (iii) and (iv)). Clearly such an incentive to merge may create non competitive effects and increase the risk.

However it is still important to discuss some conditions that need to be to imposed on a risk measure prior to implementation. For example, such a condition can be the requirement that the role of any new risk measure is to reduce risk. This condition is not satisfied by the Value at Risk

defined as a conditional quantile. Indeed the portfolio manager has an incentive to modify the usual mean-variance strategy and to select a portfolio allocation which minimizes the VaR under a constraint on the expected value of the portfolio [see e.g. Foellmer, Leukert (1998), Gourieroux, Laurent, Scaillet (2000)]. When applied to a portfolio which includes derivatives, this strategy implies much riskier positions on derivatives than the standard mean-variance strategy. This is easy to see; in order to diminish the probability of loss, which is the only constraint, the portfolio manager increases the size of loss. Such a strategy is prevented under the variance-based measure of risk. In general, such strategies can be avoided by imposing several constraints, such as joint constraints on the Value at Risk and the tail VaR. However the method for fixing the required capital as a function of the VaR and tail VaR is not known.

7.2   Infrequent Extreme Risks and Utility Functions

Before defining coherent risk measures it is necessary to specify the risks to investigate and to describe the agent aversions for these risks. Intuitively, we wish to study extreme risks that induce large losses, but are infrequent. Otherwise extreme risks could be examined and easily hedged for by the investors. Gourieroux, Monfort (2000) have developed a simple framework for examining this problem. Typically a sample of excess returns Y featuring infrequent extreme risks is obtained when the distribution of Y is a mixture of gaussian distributions:

$$Y \sim \alpha N[m, \frac{1}{\alpha}\Omega_1] + (1-\alpha)N[m, \frac{1}{1-\alpha}\Omega_2].$$

The mean is $EY = m$ and the variance $VY = \Omega_1 + \Omega_2$. However when α tends to zero, the first regime becomes infrequent whereas the covariance matrix $\frac{1}{\alpha}\Omega_1$ tends to infinity, creating extreme risk. It is easy to check that the standard utility functions such as the exponential (or CARA: Constant Absolute Risk Aversion) utility function, are not appropriate in the presence of infrequent extreme risks. Indeed there is zero demand for risky assets from investors who maximize an expected CARA utility function. As a consequence, there is no trading of these assets at equilibrium.

Gourieroux, Monfort (2000) characterized the class of utility functions for which there exists a nonzero demand for infrequent extreme risks. These functions may be written as:

$$U(w) = -\int (w-x)^- dG(x) + cw, \tag{7.2}$$

where G is a cumulative distribution function and c a nonnegative scalar. These functions are called LIRA for Left Integrable (absolute) Risk Aversion.

The associated expected utility has a simple expression . Indeed we get for $c = 0$:

$$E_w U(w) = -E_w E_x (w - x)^-,$$

where E_w and E_x denote the expectations with respect to the distribution of the portfolio value and to distribution G, respectively. By commuting the expectations, we get:

$$E_w U(w) = -E_x E_w (w - x)^- = -E_x P[x], \tag{7.3}$$

where $P[x]$ is the price of a European put written on W with strike X (computed under the historical probability). The expected utility is the opposite of the average price of puts, where the average is the strike average. This is an interesting interpretation of expected utility, since it links the treatment of extreme risk to the price of puts with strikes selected in an appropriate way.

7.3 The Dynamics of Infrequent Extreme Risks

Infrequent extreme risks have to be analyzed in a dynamic framework. Loosely speaking in the case of a single asset, extreme risks correspond to infrequent jumps in the return trajectory that cause large negative returns (or large positive returns, if the quantity of assets in the portfolio is negative). The following questions concerning the dynamics have been considered in the literature.

i) How to construct a dynamic model that allows for infrequent extreme risks, and is compatible with some stylized facts, such as the clustering of extreme risks and the possibility for standard and extreme risks to have very different dynamics? The approach followed by Gourieroux, Jasiak (2001) a,c is an example of this literature. They introduced Levy distributions to represent the conditional distribution of returns. The family involves four parameters: a location parameter, a scale parameter, a skewness parameter and a tail parameter. The four parameters are considered as stochastic factors with their own dynamics. Distinct dynamics of standard and extreme risks is obtained by introducing different serial dependence for the scale and tail parameters. It also allows for clustering of extreme risks, when the dynamics of the stochastic tail parameter features a unit root.

ii) Another important question concerns the misspecifications (also called model risks in the VaR framework). What arises when infrequent extreme risks exist but are neglected in a dynamic model? It has been proven that this misspecification induces a spurious long memory effect [see e.g. Lobato, Savin (1997), Diebold, Inoue (2001), Gourieroux, Jasiak (2001)b, Gourieroux, Robert (2001)]. In some sense it is a good news that the omission of the occurence of extreme risks induces serial smoothing of theValue at Risk, which can be interpreted as a kind of implicit insurance in time against stochastic risks.

7.4   Portfolio of a large number of assets

To understand why a large number of assets can cause a difficulty for risk analysis, let us review the standard mean-variance framework and assume that the conditional distribution of returns is multivariate normal $Y \sim N[m, \Sigma]$. The allocation of a mean-variance efficient portfolio is proportional to $a = \Sigma^{-1}(m - r_f e)$, where r_f is the riskfree rate and $e = (1, ..., 1)'$ [see e.g. Gourieroux, Jasiak (2001)a, section 3.4]. The joint dependence between the n assets is summarized by the volatility matrix Σ or better by its spectral decomposition. Let us consider the eigenvalues ranked in descending order $\lambda_1 > ... > \lambda_n$ and the corresponding eigenvectors $a_1, ..., a_n$, say. a_1 [resp. a_n] provides the portfolio allocation with the largest [resp. smallest] return volatility. Moreover, if $m = 0$ (the efficient market hypothesis) we see from equation (2.7) that a_1 [resp. a_n] maximizes [resp. minimizes] the gaussian Value at Risk; this result is valid for any risk level α.

Finally, when n is large, the smallest eigenvalue λ_n is close to zero. This can give a spurious impression of perfect arbitrage opportunity,, and implies a rather inaccurate computation of Σ^{-1} and of the optimal allocation. A number of methods have been proposed in the mean-variance framework to avoid misunderstanding and to correct for the lack of robustness.

Despite of much efforts, analogous approaches for handling the fat conditional tails have not been yet developed. Some important questions remain to be answered:

i) How to model tails that depend on portfolio allocation, that is are gaussian for some allocations and Pareto for others?

ii) Are the VaR minimizing (resp. maximizing) allocations independent of the risk level α? If they are dependent, what is the α dependence pattern?

iii) In the mean-variance framework, can the structure of dependence be simplified by imposing, for instance, an equicorrelation constraint? How to define a notion of equidependence in a nongaussian framework? [see e.g. Gourieroux, Monfort (2002)].

7.5   Extreme Value Theory

The analysis of stochastic properties of extremes is a significant part of probability and statistical theory [see e.g. Embrechts, McNeil, Straumann, Kaufmann (1998) for a survey oriented towards applications to insurance and finance]. The extreme value theory (EVT) has been developed for applications in various fields, initially not including finance. Nowadays, it is used to study extreme risks on large portfolios of individual contracts, and to predict the occurrence and size of the centenary wave, for example. We can distinguish different related topics treated by the EVT in the one-dimensional framework. These are:

i) the definition end estimation of the magnitude of the tail;

ii) the asymptotic behavior of the sample mean $\frac{1}{T} \sum_{t=1}^{T} y_t$, from the point of view of both the law of large numbers and large deviation properties;

iii) the asymptotic behavior of the sample maximum: $M_T = \max_{t \in \{1,...,T\}} y_t$;

iv) the distributional properties of the count process measuring the dates at which the process (y_t) is larger than a given threshold γ_T, function of T.

Due to initial domain of applications other than finance and to mathematical complexity, the EVT has essentially been developed for i.i.d. observations or for data with rather simple dynamics. A limited number of results exists for the complicated nonlinear dynamics encountered in Finance [see e.g. Hsing (1991), Resnick, Starica (1995) for estimation of a tail index, Robert(2000) for determination of the tail parameter in an α-ARCH model, or Gourieroux, Robert (2001) for complete analysis of stochastic unit root models].

Let us now discuss how these results might be used for improving the specification in the presence of extreme risks or for estimation of the Value at Risk.

i) **Magnitude of tail**

EVT provides classifications of tails compatible with the asymptotic theory. Such a classification has been given for instance in section 2.2ii). The EVT also explains how to estimate the tail parameter and describes the asymptotic properties of different estimators of the tail index. Such an estimator (i.e. the Hill estimator) has been used in one of the model building approaches [see section 3.3].

It is well-known that the Hill estimator and its extensions are not very accurate since it is difficult to estimate the characteristics of an infrequent event. A particularly serious problem is due to the fact that this type of estimators vary with the number of observations, possibly in a very erratic way. Moreover, the properties of such estimators have been established in an i.i.d. framework (whereas nonlinear serial dependence is crucial in Finance) and under the assumption that the risk level α tends to zero when the number of observations tends to infinity (whereas it is small, but fixed, according to the perception of regulators).

ii) **Asymptotic behavior of the sample mean**

These results can be applied to a sample mean $\frac{1}{h}(y_t + y_{t+1} + \ldots + y_{t+h})$ to study the dependence of the term structure of VaR on serial dependence and on the tails of the conditional distribution. An illustration is given in section 2.2 iii) for a simple case of i.i.d. α-stable distributed returns to show that the term structure depends on $h^{1/a}$, where a is the stability coefficient.

iii) **Asymptotic behavior of the maximum**

This part of the EVT is not very useful for determination of the Value at Risk for market risk. Indeed the maximum operator is not involved in the computation of the portfolio value (except in the case of a derivative written on the maximum of the return over a contractual period). The theoretical results are typically used for insuring against catastrophic events. As such they will likely be useful for defining the capital requirement for operational risk due to events such as a hacker attack aimed at a computer system of the bank, the closure of trading room due to fire [see the example of Credit Lyonnais in France], etc. But clearly this problem is more related to the domain of insurance than to finance.

iv) **The count process of large events**

This part of the theory explains how the distribution of the count process of large events depends on nonlinear serial dependence and on the tail of conditional distribution [see e.g. Gourieroux, Robert (2001) for a detailed illustration of this relation]. It is potentially useful for two important purposes which are: a) the prediction of the date (and magnitude) of the next future loss, and b) the control of VaR implemented by banks. The observed distribution of dates on which losses were recorded can be compared to the distribution of a process of exceedances.

8   Concluding Remarks

The review of literature on the Value at Risk given in this chapter emphasizes the variety of financial assets for which risk has to be measured, controlled and managed. To these assets belong assets and derivatives traded on an organized financial market or over the counter. Despite the variety of assets and the large number of techniques developed for the computation of VaR, it is possible to point out a common feature in many existing methodologies. They all rely on internal models for the underlying asset dynamics, derivative pricing and assessment of default probabilities. Also, the internal models often assume that risk is characterized by latent state variables. Therefore, the VaR needs to be determined by Monte-Carlo simulations of future prices and default evolutions.

The intense use of internal models requires strict monitoring. In this respect, the Basle Committe has explored two types of regulation. The first one is standardizing the internal models used by banks. The second one is conducting sensitivity analysis of VaR with respect to various deviations from the internal models.
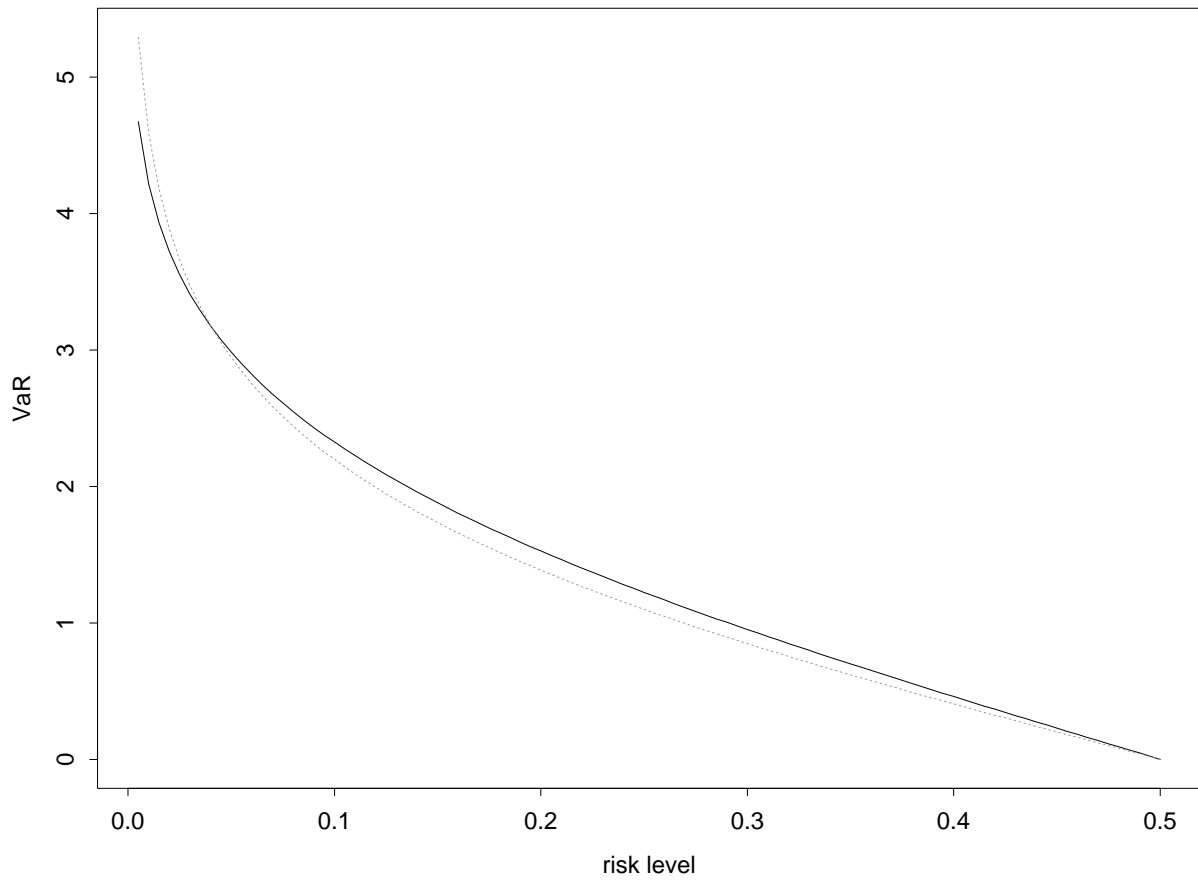
Fig 2.1: Comparison of Normal and Logistic VaR

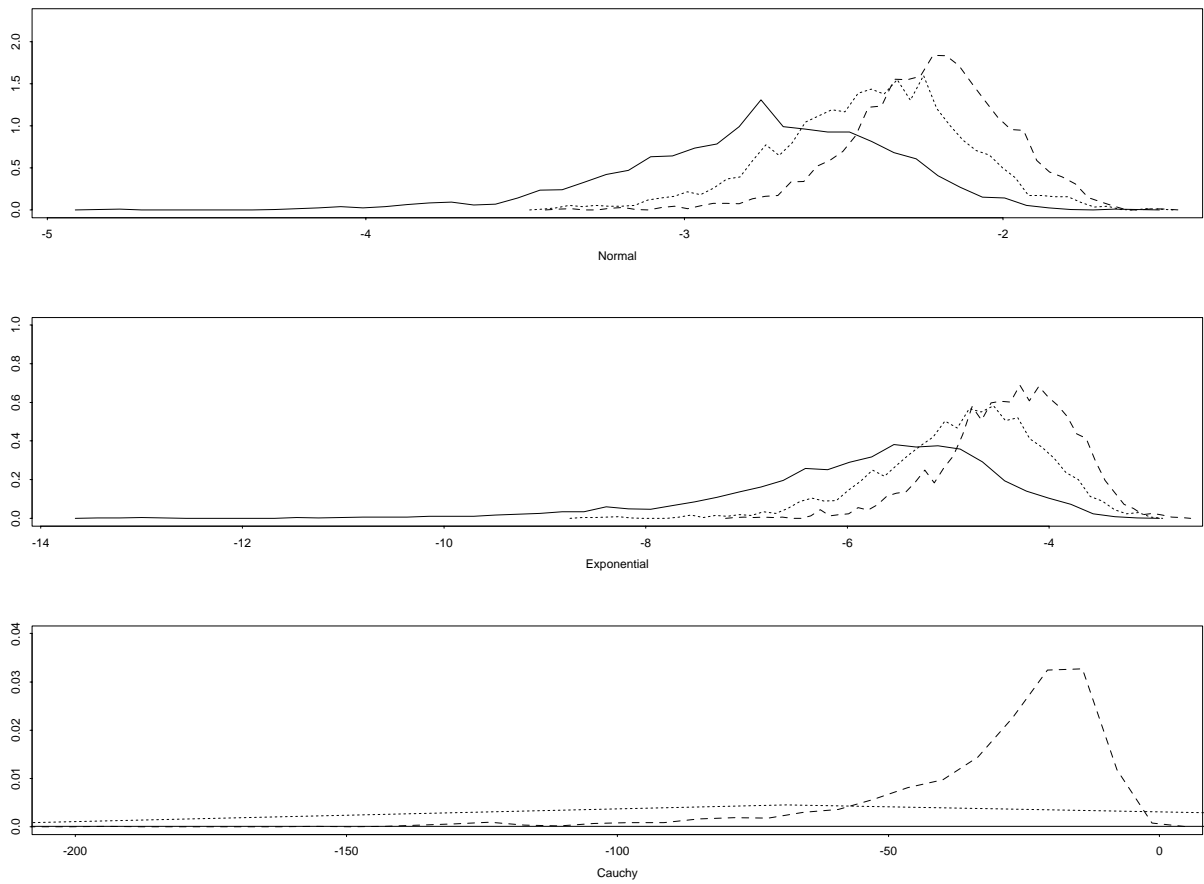Fig 3.1: Finite Sample Distributions of 1% Empirical Quantile

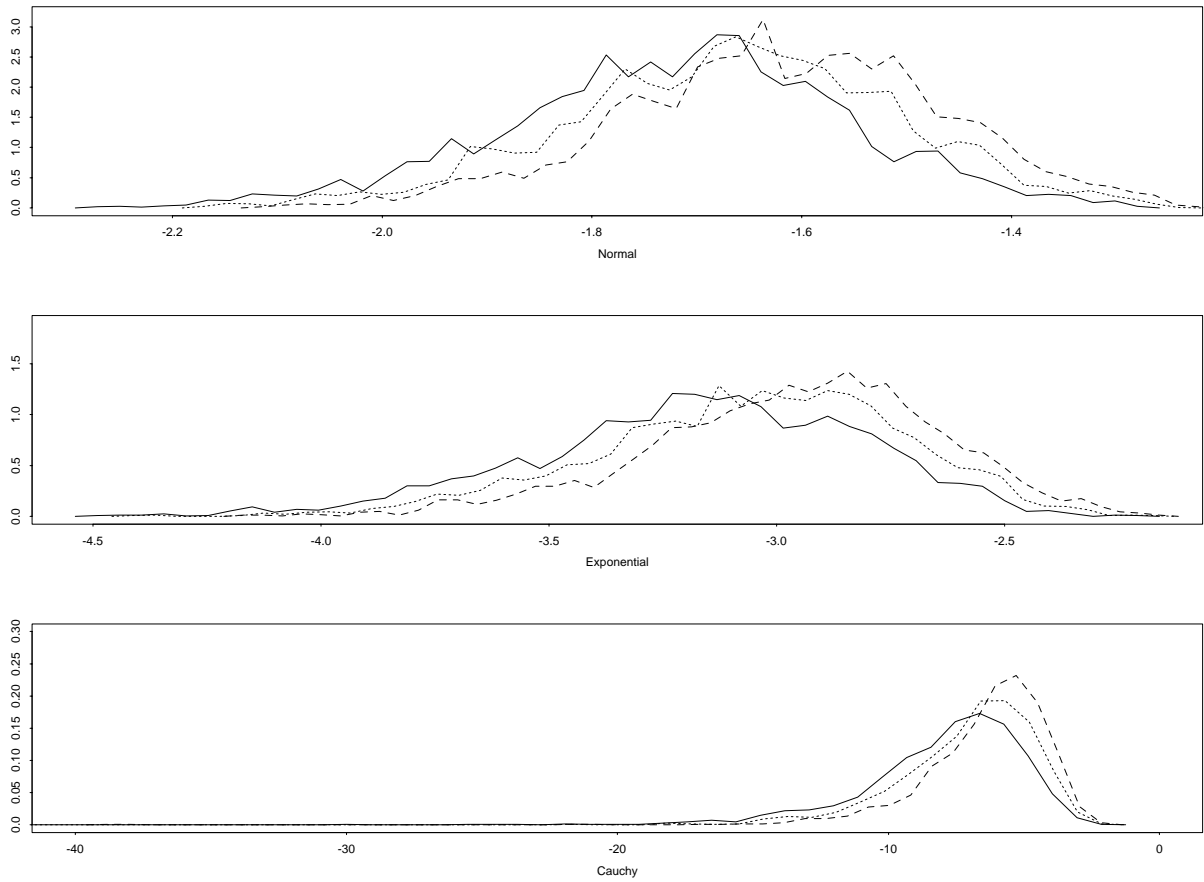## Fig 3.2: Finite Sample Distributions of 5% Empirical Quantile

Normal

Exponential

Cauchy

Fig 3.3: i.i.d. Gaussian Price Changes
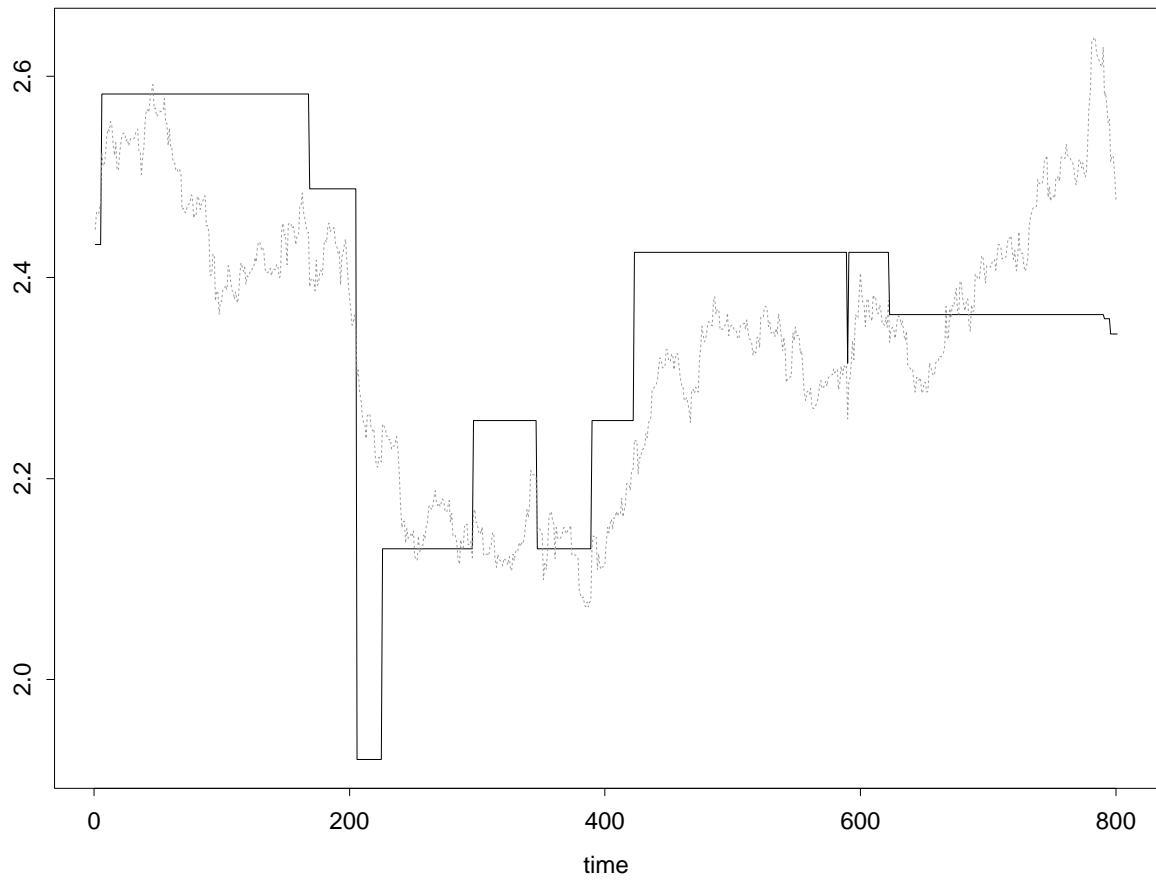

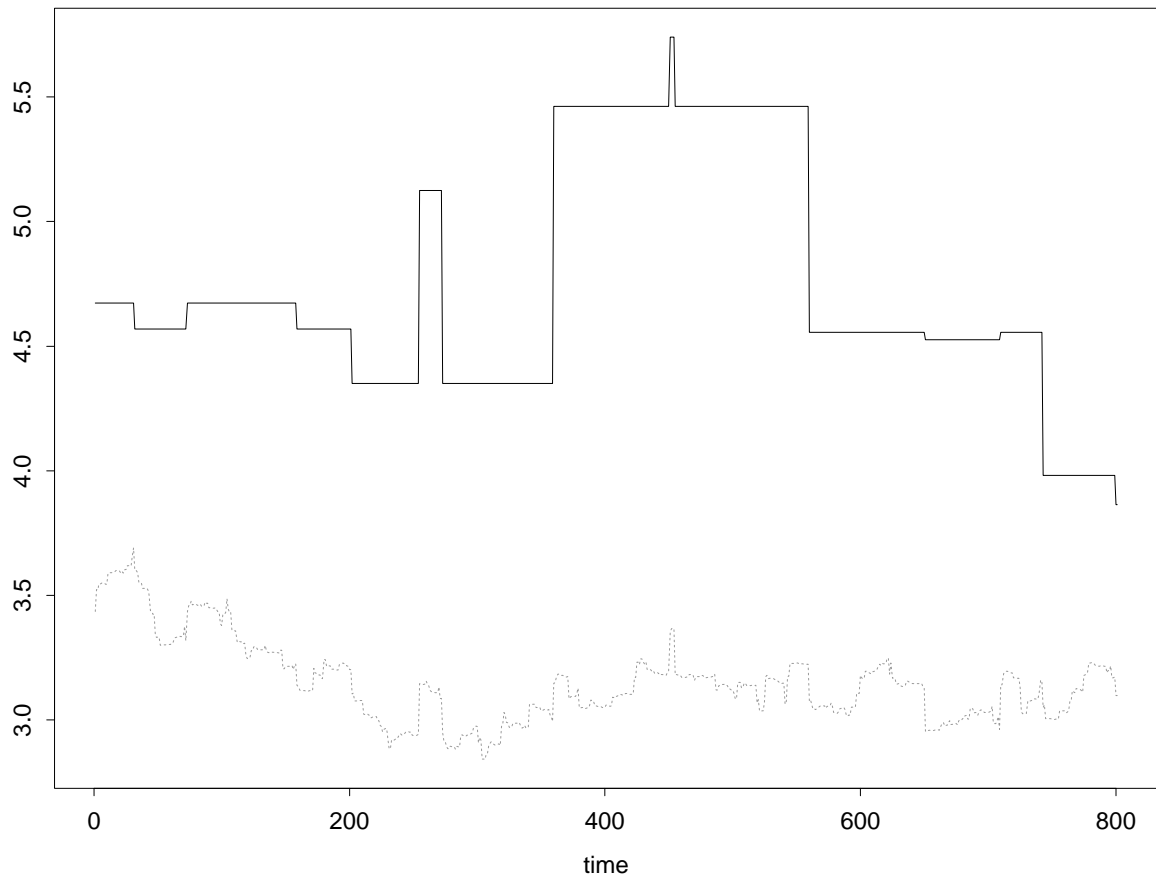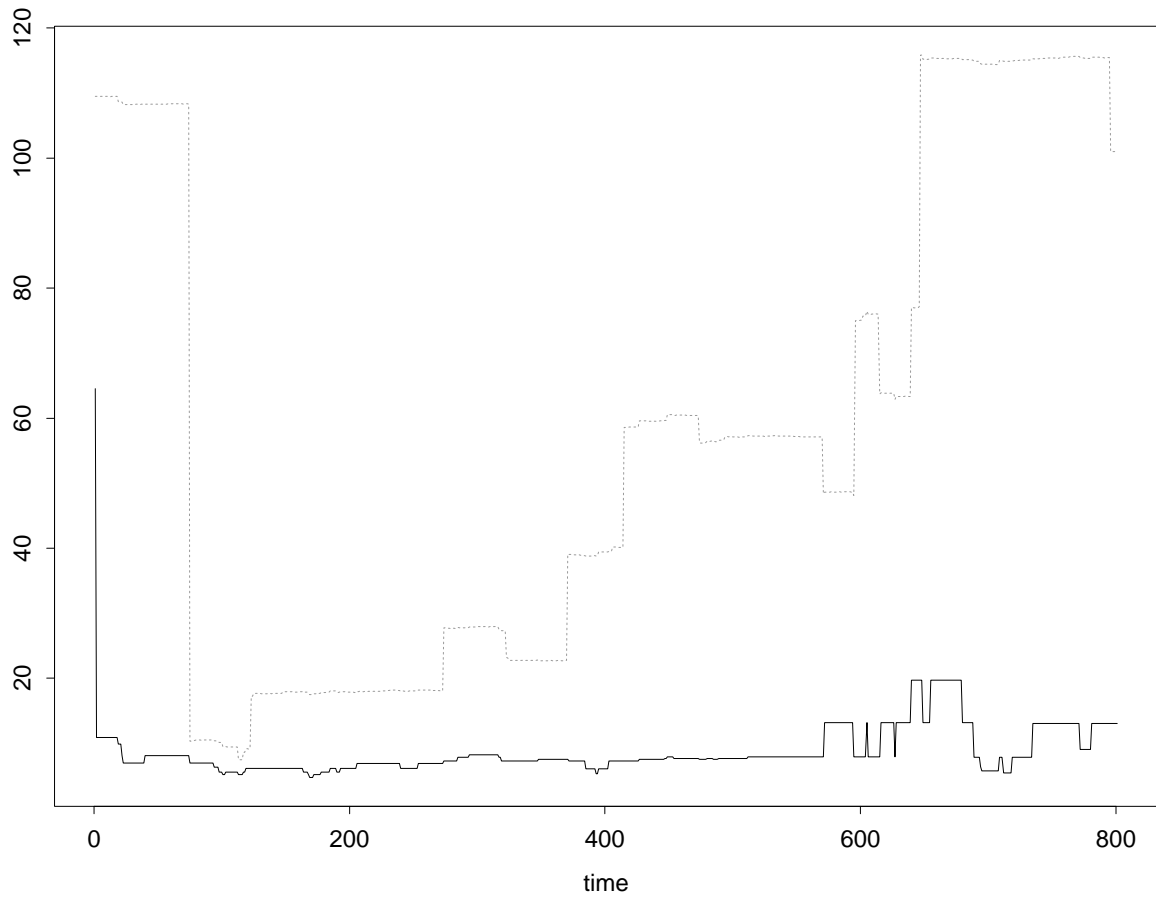time

Fig 3.4: i.i.d. Double Exponential Price Changes

Fig 3.5: i.i.d. Cauchy Price Changes

References

[1] Ait-Sahalia, Y., and A.W. Lo: Nonparametric Risk Management and Implied Risk Aversion", Journal of Econometrics, 94, 9-51.

[2] Artzner, P.,and F., Delbaen (1992) : "Credit Risk and Repayment Option", ASTIN Bulletin, 22, 81-96.

[3] Artzner, P., Delbaen, F., Eber, J. and D., Heath (1997) : "Thinking Coherently", Risk, 10, 68-71.

[4] Arzac, E., and V., Bawa (1977) : "Portfolio Choice and Equilibrium in Capital Markets with Safety First Investors", Journal of Financial Economics, 4, 277-288.

[5] Barone-Adesi, G. and K., Giannopoulos (1996): "A Simplified Approach to the Conditional Estimation of VaR", Futures and Options World, October, 68-72.

[6] Barone-Adesi, G. and K., Giannopoulos (1999)a: "The Case for Nonparametric Market Risk Measures", Risk Professional, 1, December 18-20.

[7] Barone-Adesi, G., Giannopoulos K., and L., Vosper (1999)b: "VaR without Correlations for Nonlinear Portfolios", Journal of Futures Markets, 19, 583-602.

[8] Basle Committee on Banking Supervision (1995) :"An Internal Model-Based Approach to Market Risk Capital Requirements", Basle, Switzerland.

[9] Basle Committee on Banking Supervision (1996)a "Supervisory Framework for the Use of Backtesting in Conjunction with the Internal Approach to Market Risk Capital Requirement", Basle, Switzerland.

[10] Basle Committee on Banking Supervision (1996)b : "Overview of the Amendment to the Capital Accord to Incorporate Market Risk", Basle, Switzerland.

[11] Basle Committee on Banking Supervision (1996)c : "Amendment to the Capital Accord to Incorporate Market Risk", Basle, Switzerland.

[12] Basle Committee on Banking Supervision (2000): "Principles for the Management of Credit Risk", Basle, Switzerland.

[13] Basset, G. and R., Koenker (1978): "Regression Quantiles", Econometrica, 46, 33-50.

[14] Bassi, F., Embrechts, P., and M., Kafetzaki (1997) : "Risk Management and Quantile Estimation", in Practical Guide to Heavy Tails, eds. Adler, R., Feldman, R., and M., Taqqu, Birkhauser, Boston.

[15] Beder, T. (1995) : "VaR : Seductive, but Dangerous", Financial Analyst Journal, Sep-Oct, 12-21.

[16] Beirlant, J., Broniatowski, M., Teugels, J., and P., Vynckier (1995) : "The Mean Residual Life Function at Great Age : Applications to Tails Estimation", Journal of Statistical Planning and Inference, 45, 21-48.

[17] Bielecki, T., and M., Rutkowski (2000) : "Credit Risk : Modelling, Valuation and Hedging", forthcoming Springer-Verlag, New-York.

[18] Billio, M., and L., Pelizzon (2000): "Value at Risk: A Multivariate Switching Regime Approach", Journal of Empirical Finance, 7, 531-554.

[19] Black, F., and M., Scholes (1973): "The Pricing of Options and Corporate Liabilities", Journal of Political Economy, 81, 637-659.

[20] Boos, D. (1984) : "Using Extreme Value Theory to Estimate Large Percentiles", Technometrics, 26, 33-39.

[21] Bouchaud, J.P. (2000): "Theory of Financial Risks: From Statistical Physics to Risk Management", Cambridge University Press.

[22] Chernozhukov, V. and L., Umantsev (2000): "Conditional Value at Risk: Aspects of Modeling and Estimation", mimeo, Stanford University.

[23] Coles, S. (2001) : "An Introduction to Statistical Modelling of Extreme Value", Springer Series in Statistics, Springer-Verlag, London.

[24] Crosbie, P. (1998): "Modeling Default Risk", Document 999.0000.031, KMV Corporation.

[25] Crouhy, M., Galai, D., and R., Mark (2000): "A Comparative Analysis of Current Credit Risk Models", Journal of Banking and Finance, 24, 59-117.

[26] Danielsson, J., and C., de Vries (1997) : "Tail Index and Quantile Estimation with Very High Frequency Data", *Journal of Empirical Finance*, **4**, 241 - 257.

[27] Danielsson, J., and C., De Vries (1998) : "Beyond the Sample: Extreme Quantile and Probability Estimation", London School of Economics, D.P. 298.

[28] Darolles, S., Gourieroux, C., and A., Monfort (2002) : "Corporate Debt", CREST DP.

[29] Davidson, A., and R., Smith (1990) : "Model for Exceedances over High Threshold", Journal of the Royal Statistical Society,8, 52, 393-442.

[30] Dekkers, A., Einmahl, J., and L., De Haan (1989) : "A Moment Estimator for the Index of an Extreme Value Distribution", Annals of Statistics, 17, 1833-1855.

[31] De Haan, L., Janses, D., Koedjik, K., and C. de Vries (1994) : "Safety First Portfolio Selection, Extreme Value Theory and Long Run Asset Risk", in : Extreme Value Theory and Applications, ed. Galambos, Kluwer, 471 - 487.

[32] Diebold, F. and A. Inoue (2001): "Long Memory and Regime Switching", Journal of Econometrics, 105, 131-160.

[33] Dowd, K. (1998): "Beyond Value at Risk. The New Science of Risk Management", Wiley and Sons, London.

[34] Duffee, G. (1999) : "Estimating the Price of Default Risk", The Review of Financial Studies, 12, 197-226.

[35] Duffie, D., and N., Garleanu (2001) : "Risk and Valuation of Collateralized Debt Obligations", Stanford Univ. DP.

[36] Duffie, D., and R., Kan (1996): " A Yield Factor Model of Interest Rates", Mathematical Finance, 6, 379-406.

[37] Duffie, D., and J., Pan (1997) : "An Overview of Value at Risk", Journal of Derivatives, 4, 7-49.

[38] Duffie, D., and K., Singleton (1999): "Modelling Term Structures of Defaultable Bonds", Review of Financial Studies, 12, 197-226.

[39] El Jahel, L., Perraudin, W., and P., Sellin (1998) : "Value at Risk for Derivatives", DP IFR 50, Birbeck College.

[40] Embrechts, P., Kluppelberg, G., and T., Mikosch (1997) : "Modelling Extremal Events", Springer, Applications of Mathematics.

[41] Embrechts, P., Mc Neil, A., Straumann, D., and R. Kaufmann (1999): "Modeling Extremal Events for Insurance and Finance", 15th International Summer School of the Swiss Association of Actuaries.

[42] Embrechts, P., Resnick, S., and G., Samorodnitsky (1998) : "Living on the Edge", Risk, 11, 96-100.

[43] Engle, R. (1982) : "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation", *Econometrica*, **50**, 987 - 1008.

[44] Engle, R., and S., Manganelli (1999): "CaViaR: Conditional Autoregressive Value at Risk by Regression Quantiles", D.P. University of California at San Diego.

[45] Estrella, A. (1996): "Taylor, Black and Scholes: Series Approximations and Risk Management Pitfalls", In Risk Measurement and Systematic Risk, Proceedings of the Joint Central Bank Research Conference, Board of Governors of the Federal Reserve System (ed.), Washington D.C. 359-379.

[46] Falk, M. (1985) : "Asymptotic Normality of the Kernel Quantile Estimator", The Annals of Statistics, 13, 428-433.

[47] Figlewski, S., and L., White (1995): "Orange County: Don't Blame Derivatives", Stern Business, New York University, Spring, 33-41.

[48] Finger, C. (1997): "When is a Portfolio of Options Normally Distributed", Quarterly Risk-Metrics Monitor, Third Quarter, 33-41.

[49] Foellmer, H., and P., Leukert (1998) : "Quantile Hedging", forthcoming Finance and Stochastics.

[50] Fons, J. (1994): "Using Default Rates to Model the Term Structure of Credit Risk", Financial Analyst Journal, Sept.-Oct., 25-32.

[51] Garman, M. (1996) : "Improving on VaR", Risk, 9, 61-63.

[52] Garman, M. (1997) : "Taking VaR to Pieces", Risk, 10, 70-71.

[53] Genest, C., and L., Rivest (1993) : "Statistical Inference Procedures for Bivariate Archimedean Copulas", Journal of the American Statistical Association.

[54] Gizycki, M., and N., Hereford (1995): "Assessing the Dispersion in Banks Estimates of Market Risk: The Results of a Value at Risk Survey", Discussion Paper, October, Australian Prudential Regulation Authority.

[55] Golub, B., and L., Tilman (1997): "Measuring Yield Curve Risk Using Principal Component Analysis, Value at Risk and Key Durations", The Journal of Portfolio Management, 23, 72-84.

[56] Gourieroux, C., and J., Jasiak (1999)b : "Nonlinear Innovations and Impulse Response", CEPREMAP DP 9906.

[57] Gourieroux, C., and J., Jasiak (2000): "Local Likelihood Density Estimation and Value at Risk", CREST DP.

[58] Gourieroux, C., and J., Jasiak (2001)a: "Econometrics of Finance", Princeton University Press, Princeton.

[59] Gourieroux, C., and J., Jasiak (2001)b: "Memory and Infrequent Breaks", Economics Letters, 70, 1, 29-41.

[60] Gourieroux, C., and J., Jasiak (2001)c: "Dynamic Factor Models", Econometric Reviews, 20, 4, 385-424.

[61] Gourieroux, C., Laurent J.P., and O., Scaillet (2000) : "Sensitivity Analysis of Values at Risk", Journal of Empirical Finance, vol 7, 225-247.

[62] Gourieroux, C., and A., Monfort (1998) : "Statistics and Econometric Models", Vol I, Cambridge University Press.

[63] Gourieroux, C., and A., Monfort (2000): "Infrequent Extreme Risks", CREST, D.P.

[64] Gourieroux, C., and A., Monfort (2002): "Equidependence in Qualitative and Duration Models", CREST D.P.

[65] Gourieroux, C., and C. Robert (2001): "Stochastic Unit Root Models", CREST, D.P.

[66] Group of Thirty, The (1993): "Derivatives: Practices and Principles", Global Derivatives Study Group, G30, Report, July.

[67] Hamilton, J. (1989): "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle", Econometrica, 57, 357-384.

[68] Hill, D. (1975) : "A Simple Approach to Inference About the Tail of a Distribution", The Annals of Statistics, 13, 331-341.

[69] Horowitz, J. (1992): "A Smoothed Maximum Score Estimator for the Binary Response Model", Econometrica, 60, 505-531.

[70] Hsing, T. (1991) : "On Tail Index Estimation Using Dependent Data", Annals of Statistics, 19, 1547-1569.

[71] Huisman, R. and alii (1997): "VaR+: Fat Tails in Financial Risk Managament", Journal of Risk, 1, 47-61.

[72] Hull, J., and A., White (1987): "The Pricing of Options on Assets with Stochastic Volatilities", Journal of Finance, 42, 281-300.

[73] Hull, J. and A. White (1992): "The Price of Default", Risk, 5, 101-103.

[74] Hull, J., and A., White (1998)a: "Incorporating Volatility Updating into the Historical Simulation Method for Calculating Value at Risk", paper presented at the Fields Institute, Toronto.

[75] Hull, J., and A., White (1998)b: "Value at Risk when Daily Changes in Market Variables are not Normally Distributed", Journal of Derivatives, 5, 9-19.

[76] Hull, J., and A., White (2001) : "Valuing Credit Default Swaps II : Modelling Default Correlations", Journal of Derivatives, 8, 12-23.

[77] Jackson, P., Mande, D., and W. Perraudin (1997): "Bank Capital and Value at Risk", Journal of Derivatives, Spring 73-89.

[78] Jamshidian, F., and Y., Zhu (1997): "Scenario Simulations: Theory and Methodology", Finance and Stochastics, 1, 43-67.

[79] Janosi, T., Jarrow, R., and Y., Yildirim (2001) : "Estimating Default Probabilities in Equity Prices", Cornell Univ. DP.

[80] Jansen, D., Koedijk, K., and C., de Vries (1998) : "Portfolio Selection with Limited Downside Risk", Maastricht Univ.

[81] Jarrow, R (2000) : "Default Parameter Estimation Using Market Prices", forthcoming, Financial Analysts Journal.

[82] Jarrow, R, Lando, D., and S. Turnbull (1997): " A Markov Model for the Term Structure of Credit Risk Spreads", Review of Financial Studies, 10, 481-523.

[83] Jarrow, R., Lando, D. and F., Yu (2001) : "Default Risk and Diversification : Theory and Applications", Cornell Univ-DP.

[84] Jarrow, R., and S., Turnbull (2000) : "The Intersection of Market and Credit Risk", Journal of Banking and Finance, 24, 271-299.

[85] Joe, H. (1997): "Multivariate Models and Dependence Concepts", Chapman and Hall.

[86] Jorion, P. (1997): "Value at Risk: The New Benchmark for Controlling Market Risk", Chicago, McGraw-Hill.

[87] Kliger, D., and O., Sarig (2000) : "The Information Value of Bond Ratings", Journal of Finance, 55, 2879-2902.

[88] Kotz, S., and S., Nadarajah (2000) : "Extreme Value Distributions", Imperial College Press, London.

[89] Lando, D. (1997): "Modelling Bonds and Derivatives with Default Risk", in M. Dempster and S., Pliska ed., Mathematics of Derivative Securities, Cambridge University Press, 369-393.

[90] Lando, D. (1998): "On Cox Processes and Credit Risky Securities", Review of Derivatives Research, 2, 99-120.

[91] Lando, D. (2000): "Some Elements of Rating Based Credit Risk Modeling" in N., Jegadeesh and B., Tuckman ed., Advanced Fixed Income Valuation Tools, Wiley, 193-215.

[92] Lando, D. and T., Skoderberg (1998): "Analysing Rating Transition and Rating Drift with Continuous Observations", forthcoming The Journal of Banking and Finance.

[93] Levy, H., and M. Sarnat (1972) : "Safety First. An Expected Utility Principle", Journal of Financial and Quantitative Analysis, 7, 1829-1834.

[94] Lobato I., and Savin N. (1997): "Real and Spurious Long Memory Properties of Stock Market Data", Journal of Business and Economic Statistics, 16, 261-283.

[95] Longerstay, J. (1996): "An Improved Methodology for Measuring VaR", RiskMetrics Monitor, Second Quarter, J.P. Morgan.

[96] Longin, F. (1997) : "Beyond the VaR", DP 97-011, ESSEC.

[97] Longin, F. (2000): "From Value at Risk to Stress Testing: The Extreme Value Approach", Journal of Banking and Finance, 7, 1097 - 1130.

[98] Longstaff, F., and E. Schwartz (1995): "A Simple Approach to the Valuing Risky Fixed and Floating Rate Debt", Journal of Finance, 50, 789-819.

[99] Malz, A. (2000): " Vega Risk and the Smile", Working Paper, The RiskMetrics Group.

[100] Mandelbrot, B. (1963) : "The Variation of Certain Speculative Prices", *Journal of Business*, **36**, 394 - 419.

[101] Marshall, C. and M., Siegel (1997): "Value at Risk: Implementing a Risk Measurement Standard", The Journal of Derivatives, 4, 91-110.

[102] Merton, R. (1974): "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates", The Journal of Finance, 29, 449-470.

[103] Mina, J., and A., Ulmer (1999): "Delta-Gamma Four Ways", Working Paper, the RiskMetrics Group.

[104] Minnich, M. (1998): "A Primer on Value at Risk", in Fabozzi, Frank, J. (ed.), Perspectives on Interest Rate Risk Management for Money Managers and Traders, chapter 3, 39-50.

[105] Moody's Investors Service (1997): " Moody's Rating Migration and Credit Quality Correlation, 1920 -1996", July.

[106] Morgan Guaranty Trust Company (1996): "RiskMetrics Technical Document", 4th ed.

[107] Nielsen, L., Saa-Requejo, J., and P., Santa-Clara (1993): "Default Risk and Interest Rate Risk: The Term Structure of Default Spreads", INSEAD.

[108] Oda, N., and J., Muranoga (1997): "A New Framework for Measuring the Credit Risk of a Portfolio - Ex-VaR Model", Discussion Paper, Institute of Monetary and Economic Studies, Bank of Japan, 97, 1-45.

[109] Phelan, M. (1995): Probability and Statistics Applied to the Practice of Financial Risk Management: The Case of J.P. Morgan's RiskMetrics", Working Paper, Wharton School of Management.

[110] Pickands, J. (1975) : "Statistical Inference Using Order Statistics", Annals of Statistics, 3, 119-131.

[111] Pritsker, M. (1997): "Evaluating Value at Risk Methodologies: Accuracy versus Computational Time", Journal of Financial Services Research, October, 201-241.

[112] Resnick, S. (1987) : "Extreme Values, Regular Variation and Point Processes", Springer, New-York.

[113] Resnick, S., and C., Starica (1995) : "Consistency of Hill's Estimator for Dependent Data", Journal of Applied Probability, 32, 139-167.

[114] Ridder, T. (1997) : "Basics of Statistical VaR- Estimation", SGZ-Bank, Frankfurt/Karlsruhe.

[115] Robert, C. (2000) : "Extremes of α-ARCH Models", in Measuring Risk in Complex Stochastic Systems, Lecture Notes in Statistics, Springer, Vol 147, Ed. Franke, Hardle, Stahl.

[116] Schachter, B., and J., Butler (1998): "Estimating Value at Risk with a Precision Measured by Combining Kernel Estimation with Historical Simulation", Review of Derivatives Research, 1, 371-390.

[117] Schonbucher, P. (2000): "Factor Models for Portfolio of Credit Risk", Bonn Univ. DP.

[118] Sklar, A. (1959) : "Fonctions de répartition à n dimensions et leurs marges", Publ. Inst. Stat. Univ. Paris, 8, 229-231.

[119] Smith, R. (1989) : "Extreme Value Analysis of Environmental Time Series : An Application to Trend Detection in Ground Level Ozone", Statistical Science, 4, 367-393.

[120] Smithson, C. (1996)a : "Value at Risk ", Risk, 9, 25-27.

[121] Smithson, C. (1996)b : "Value at Risk (2)", Risk, 9, 38-39.

[122] Venkataraman, S. (1997): "Value at Risk for a Mixture of Normal Distributions. The Use of Quasi Bayesian Estimation Techniques", Economic Perspectives, Federal Reserve Bank of Chicago, 21, 2-13.

[123] Wilson, T. (1996) : "Calculating Risk Capital", in Alexander, Carol eds, The Handbook of Risk Management and Analysis, Chichester, Wiley, 193-232.

[124] Zangari, P. (1996): "An Improved Methodology for Measuring VaR", Quarterly RiskMetrics Monitor, Second Quarter, 7-25.