A Gentle Introduction to Value at Risk

Laura Ballotta[1]
Gianluca Fusai[2]

March 28, 2017

[1]Senior Lecturer in Finance, Faculty of Finance, Cass Business School, City University London, 106 Bunhill Row, London EC1Y 8TZ, UK. Email: l.ballotta@city.ac.uk.

[2]Full Professor in Financial Mathematics, Department SEI, U. Studi del Piemonte Orientale, Via Perrone 18, 28100 Novara (Italy) and Reader in Finance, Cass Business School, City University London, 106 Bunhill Row, London EC1Y 8TZ, UK. Email: gianluca.fusai@uniupo.it, gianluca.fusai.1@city.ac.uk.

Contents

Foreword

The idea of this document is to provide the reader with an intuitive introduction to the main tools in risk analysis. Large part of this material has been used in MSc courses (30 hours) in quantitative finance on risk analysis at Università del Piemonte Orientale and Cass Business School.

We would like to emphasize that this document is very much work in progress and we would like to encourage readers to get in touch with us with feedback, comments, suggestions for additions and, of course, corrections of typos. All of these will be gratefully acknowledged in the future releases of this document.

A companion *A Gentle Introduction to Default Risk and Counterparty Credit Modelling* by Ballotta, Fusai and Marena [4] is also available on the SSRN web site.

Laura Ballotta and Gianluca Fusai

This Version: 1.1

1 Introduction

Uncertainty about commodity prices, foreign markets, new technologies, government policies, and even weather conditions can affect significantly firm earnings, so that a sound risk management plays an important role in many business decisions. This is important not only for banks, for which quite stringent rules are already at work, but also for corporate firms. Academic finance literature identifies at least three relevant issues associated with unpredicted earnings volatility: (1) higher expected costs of financial distress; (2) higher expected payments to corporate stakeholders (including higher rates of return required by owners of closely-held firms); and (3) lower tax payments due to reduction in the fluctuations of taxable income through suitable risk management policies[1]. From the management point of view, the decision about which risks to keep and which to hedge requires a comprehensive risk-audit review, see Stulz [25]. The aim of this chapter is to present a modern approach to the risk measurement of a financial position through statistical techniques which allow to describe the profit and loss (henceforth P&L) distribution of the firm's portfolio over some predetermined horizon. In particular, market practice has nowadays adopted Value at Risk (VaR) as standard risk measure[2]. Very often, this measure is supported by stress testing analysis aimed at understanding the effect of extreme movements in market variables on the portfolio's firm. For more details on the main uses of VaR at corporate level see Jorion [22].

This paper is an introduction devoted to the measurement of market risk in financial markets, with examples mainly drawn from commodity markets. In particular, we present the concept of VaR, its limits, the problems related to its estimation and backetesting. This is done at the single asset and at the portfolio level. Issues related to the measurement of portfolio risk contribution and how to cope with derivative positions are also considered. Other important issues like liquidity, operational and credit risk will not be dealt here. For a gentle introduction to the measurement of counterparty credit risk see the companion paper by Ballotta, Fusai and Marena [4].

---

[1]For a detailed discussion on these themes, see Stultz [25].

[2]VaR was introduced by J.P. Morgan to monitor the exposure created to financial institutions by their trading activities. For this reason they set up the RiskMetrics group that soon proposed VaR as a benchmark risk measure.

2 Value at Risk (VaR)

The starting point to VaR computation is the probability distribution of the P&L. If estimated properly, the P&L distribution reflects the netting and diversification effects and can be compared across portfolios.

The idea behind VaR is to build the P&L probability distribution of the bank or corporate portfolio at a given time horizon and compute the worst case loss at a given confidence level.

In order to do this, first we need to understand how our portfolio is affected by unexpected shocks in risk factors. Hence we have to decide which quantity to model: prices or returns. The advantage of using returns is that in this case the stationarity assumption, very needed to perform statistical estimates, is more realistic than in the case of prices. In the following, we let $P(t)$ be the t-value of the corporate portfolio. We also define T to be the corporate horizon, so that the P&L over the period $[t, T]$ is defined as

$$P\&L(t, T) = P(T) - P(t).$$

A related quantity is the so called holding period log-return, defined as

$$r(t, T) = \ln \frac{P(T)}{P(t)}.$$

For a variety of reason, we prefer to work with logarithmic returns rather than arithmetic (simple, linear) ones[3].

Notice that there is also an inverse mapping between prices and log-returns through the relationship

$$P(T) = P(t)e^{r(t,T)},$$

so that any modelling assumption on returns will imply a corresponding model for prices and viceversa. For example, assuming that the log-returns distribution is Gaussian implies that prices have a log-normal distribution.

In the following we introduce the standard definition of VaR.

---

[3]The main advantage of using log-returns versus simple returns is that log-returns satisfy the additivity rule, i.e. the cumulative log-return over 1 week, say, is given by the sum of five 1 day log-returns. In addition, a common assumption in finance is that returns are distributed according to a Gaussian distribution. Log-returns can take values over the real line, i.e. the same support of the Gaussian distribution, whilst linear returns are defined on the region $(-1, +\infty)$.

Definition 1 (Value at Risk) *VaR in monetary terms[4] is the maximum loss over a target horizon such that the probability that the actual loss is larger is equal to $1 - \alpha$, where α is the confidence level, i.e.*

$$\Pr\left(P(T) - P(t) < -VaR_\alpha^{P\&L}(t,T)\right) = 1 - \alpha. \tag{1}$$

We can also define the VaR using the cumulative distribution function (CDF) of log-return

$$\Pr\left(r(t,T) < -VaR_\alpha^r(t,T)\right) = 1 - \alpha.$$

Given that $P(T) = P(t)\,e^{r(t,T)}$, the P&L $VaR^{P\&L}$ and the return VaR^r are related by

$$VaR_\alpha^r(t,T) = -\ln\left(1 - \frac{VaR_\alpha^{P\&L}(t,T)}{P(t)}\right)$$

$$VaR_\alpha^{P\&L}(t,T) = P(t) \times \left(1 - e^{-VaR_\alpha^r(t,T)}\right).$$

An approximated transformation is

$$VaR_\alpha^{P\&L}(t,T) \sim P(t) \times VaR_\alpha^r(t,T).$$

Example 2 (Meaning of VaR (Dollar terms)) *The 10 day VaR at the confidence level of 90% is 5ml USD means that*

- *with probability 90%, the bank will not loose more than 5ml USD over a 10 days period,*

- *with probability 10%, the bank will loose more than 5ml USD over a 10 days period.*

Example 3 (Meaning of VaR (returns terms)) *The 1 day VaR at the confidence level of 99% is 3% means that*

- *with probability 99%, the bank will realize a return higher than -3% over a 1 day period,*

- *with probability 1%, the bank will realize a return lower than -3% over a 1 day period.*

---

[4]In the following, we use the notation $VaR^{P\&L}$ and VaR^r or simply VaR to distinguish between VaR in monetary terms and VaR in return terms.

Given the VaR and the current portfolio value of 1ml of USD, the portfolio VaR (expressed as amount of USD) is

$$VaR_{\alpha}^{P\&L}\left(t, t + \frac{1}{250}\right) = 1 \times \left(1 - e^{-0.03}\right) ml = 29,554\$,$$

i.e. we will loose less than 29,554\$ in 1 day with a 99% probability. There is still a 1% probability to loose more than 29,554\$.

In Figure 1 we plot the cumulative distribution function (CDF) of the P&L. If the CDF is continuous and strictly increasing, the VaR number is found as the intersection between the CDF and the horizontal line at the height $1 - \alpha$. Given that the CDF (at least for the case of a continuous random variable) is nothing else than the cumulative integral of the probability density function (PDF), the VaR number can also be found by looking for the abscissa such that the area under the PDF is equal to $1 - \alpha$. This is illustrated in Figure 2. Therefore, if we define F_r and f_r to be respectively the cumulative and the density probability functions of log-returns, the VaR can be defines as the number that solves one of the two following equations

$$F_r(-VaR_{\alpha}) = 1 - \alpha,$$

and

$$\int_{-\infty}^{-VaR_{\alpha}} f_r(x)dx = 1 - \alpha.$$

Example 4 (Estimating VaR for a continuous r.v.) *Let us suppose that the P&L of an interest rate swap is assumed to be uniformly distributed on the interval -1 ml, 1 ml. Then the probability density function (pdf) is*

$$f_{P\&L}(x) = \frac{1}{2},$$

if $-1 < x < 1$ and 0 elsewhere. The **cumulative** *density function is*

$$F_{P\&L}(x) = \Pr(P\&L \leq x) = \begin{cases} 0 & if \quad x \leq -1, \\ \frac{x-(-1)}{2} & if \quad -1 < x \leq 1, \\ 1 & if \quad x > 1. \end{cases}$$

The α-VaR solves

$$\Pr(P\&L \leq -VaR_{\alpha}) = 1 - \alpha,$$
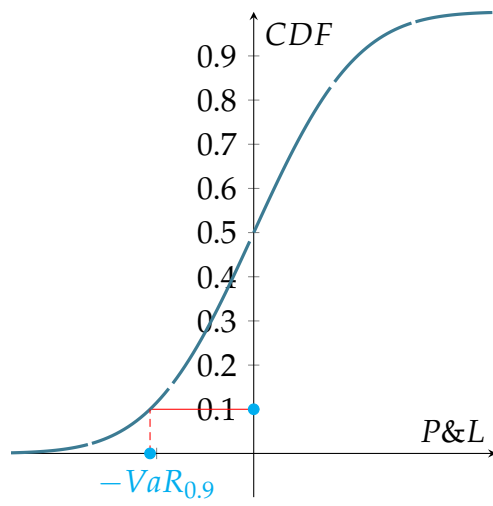
Figure 1: For a continuous r.v., the VaR is determined by the abscissa at which the CDF equals the level $1 - \alpha$.
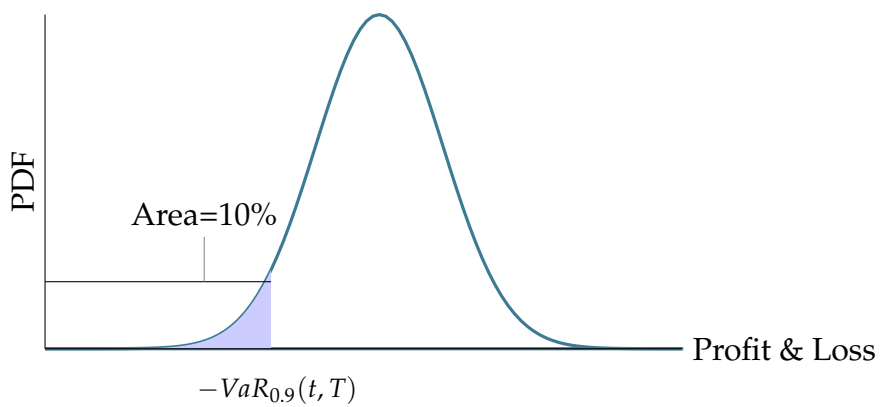
Figure 2: For a continuous r.v., the VaR is determined by the abscissa at which the area below the PDF equals the level $1 - \alpha$.

i.e.

$$\frac{-VaR_\alpha - (-1)}{2} = 1 - \alpha$$

and therefore

$$-VaR_\alpha = 2 \times (1 - \alpha) - 1,$$

i.e.

$$VaR = 2\alpha - 1.$$

The 95% VaR is $VaR_{0.95} = 0.9ml$ USD.

For any discrete distribution, the VaR is instead defined as follows

$$VaR_\alpha (t, T) = - \inf_x \{x \in \Re, \Pr (r (t, T) \le x) \ge 1 - \alpha\}.$$

i.e. it corresponds to the smallest value of the P&L such that the CDF is above (or equal) to the level $1 - \alpha$.

Example 5 (VaR of a discrete r.v.) *Let us compute the 85% VaR for a discrete r.v. having distribution as in Table 1. Therefore, we have that*

$$\inf_x \{x \in \Re, \Pr (r (t, T) \le x) \ge 0.15\} = -0.5ml \ \$,$$

so that

$$VaR_{0.85} (t, T) = 0.5ml \ \$.$$

This is illustrated in figure 3.

P&L (Ml $)	PDF	CDF	Is $CDF \ge 0.15$?
-2	0.04	0.04	no
-1	0.05	0.09	no
-0.5	0.2	0.29	yes
0	0.3	0.59	yes
0.5	0.2	0.79	yes
1	0.12	0.91	yes
1.5	0.07	0.98	yes
2	0.02	1	yes

Table 1: For a discrete r.v. the VaR is obtained by searching the smallest value at which the CDF is greater than $1 - \alpha$. In this example, $\alpha = 0.85$.
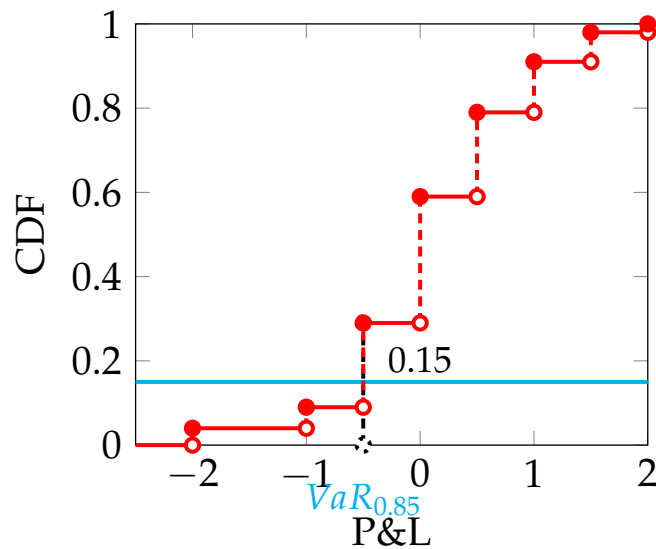
Figure 3: Computation of the VaR for a discrete r.v.

Three remarks

1. VaR can be against portfolio diversification.

2. The VaR does not describe the maximum loss.

3. VaR does not describe the losses in the left tail.

Remark 1 Given two portfolios P&L described by the random variables X and Y, we would expect to have:

$$VaR^{X+Y} < VaR^X + VaR^Y,$$

i.e. combining risks is less risky than treating the risks separately. This means that there has to be some gain from diversification. However, it is possible to present examples for which

$$VaR^{X+Y} > VaR^X + VaR^Y,$$

i.e. we say that the VaR measure is violating the subadditivity property.

Example 6 (VaR is not subadditive) *Let us consider the following example taken from Acerbi et al. 2008, w.p. Consider bonds A and B with non overlapping default probabilities. This means that if one defaults the other will not and vice*

Event	A	B	A+B	Prob.
A looses 30	70	100	170	3%
A looses 10	90	100	190	2%
B looses 30	100	70	170	3%
B looses 10	100	90	190	2%
No Loss	100	100	200	90%

Table 2: Payoffs of the portfolio invested in the two defaultable bonds.

	Prob. Distribution of A and B			Prob. Distribution of A+B		
P&L	-28.9	-8.9	1.1	-27.8	-7.8	2.2
PDF	3%	2%	95%	6%	4%	90%
CDF	3%	**5%**	100%	**6%**	10%	100%
	$VaR_{0.95}^A = VaR_{0.95}^B = 8.9$			$VaR_{0.95}^{A+B} = 27.8$		

Table 3: Payoffs of the portfolio invested in the two defaultable bonds.

versa. The two bonds have two different default states each with recovery values at 70 and 90 and probabilities 3% and 2% respectively. Otherwise they will redeem at 100. The possibile states of the world and the corresponding probabilities are given in table
The market price of the two bonds is the expected value of their payoffs

$$0.03 \times 70 + 0.02 \times 90 + 0.95 \times 100 = 98.9.$$

The portfolio is worth

$$98.9 \times 2 = 197.8.$$

The VaR computation and violation of subadditivity is shown in table 3. In particular, the portfolio VaR (27.8\$) is greater than the sum of the individual VaR (17.8\$). Therefore, in this example, the VaR is not subadditive.

Artzner et al. [2] and [3] introduced the concept of coherence for any risk metric that satisfies four axioms, including the diversification (or subadditivity) one. VaR turns out to be coherent only under special assumptions on the distribution of returns. For example, the coherence of VaR occurs when returns are normally distributed, see example 12 below, or, more in general, when they have an elliptical distribution. Artzner et al. [2] also show that a possible coherent risk metrics is the expected shortfall (or conditonal VaR or conditional tail expectation). Expected Shortfall (ES) averages the P&L scenario returns beyond the VaR threshold, so it provides additional information as to how severe the losses can be.

Definition 7 (Expected Shortfall) *ES in monetary terms is the average of the* $(1 - \alpha) \%$ *worst losses, i.e.*

$$ES_\alpha^{P\&L}(t, T) = -\mathbb{E}\left(P\&L(t, T) \Big| P\&L(t, T) \leq -VaR_\alpha^{P\&L}(t, T)\right).$$

ES can be defined in log-returns terms

$$ES_\alpha^r(t, T) = -\mathbb{E}\left(r(t, T) | r(t, T) \leq -VaR_\alpha^r(t, T)\right).$$

Given that the ES is the average of the worst losses, whilst the VaR is the best of the worst losses, we always have

$$ES_\alpha(t, T) \geq VaR_\alpha(t, T),$$

Both quantities are increasing functions of α (i.e. the higher the confidence level, the higher the exposure). In general, the difference between ES and VaR depends on how thick is the left tail of the relevant distribution

Fact 8 (Expected Shortfall for continuous random variables) *If log-returns admit a continous probability density function f_r, the Expected Shortfall can be computed according to*

$$ES_\alpha^r(t, T) = -\frac{\int_{-\infty}^{-VaR_\alpha^r(t,T)} x f_r(x) \, dx}{\underbrace{\int_{-\infty}^{-VaR_\alpha^r(t,T)} f_r(x) \, dx}_{\Pr(r(t,T) < -VaR_\alpha^r(t,T)) = 1 - \alpha}} = -\frac{1}{1 - \alpha} \int_{-\infty}^{-VaR_\alpha^r(t,T)} x f_r(x) \, dx.$$

(2)

If we are interested in the ES expressed in monetary terms, we have to compute

$$ES_\alpha^{P\&L}(t, T) = -\frac{1}{1 - \alpha} \int_{-\infty}^{-VaR_\alpha^{P\&L}(t,T)} x f_{P\&L}(x) \, dx.$$

(3)

Notice that, differently from VaR, there is no simple rule to transform the ES defined in log-return term in the ES defined in monetary terms. An approximated transformation is given by

$$ES_\alpha^{P\&L}(t, T) \sim P(t) \times ES_\alpha^r(t, T).$$

Example 9 (Computing the Expected Shortfall) *Consider a swap position. The initial value of the swap is 0. The 1 month swap P&L is assumed to be*

uniformly distributed in the range [-1, 1] million USD. The PDF of the P&L is given by

$$f_{P\&L}(x) = \begin{cases} 0 & if\, x \leq -1 \\ \frac{1}{2} & if -1 \leq x \leq 1 \\ 0 & if\, x > 1 \end{cases}$$

Therefore the CDF is given by

$$F_{P\&L}(x) = \begin{cases} 0 & if\, x \leq -1 \\ \frac{x+1}{2} & if -1 \leq x \leq 1 \\ 1 & if\, x > 1 \end{cases}$$

and the VaR is solution of the equation

$$\frac{-VaR + 1}{2} = 1 - \alpha,$$

so that

$$VaR_{\alpha}^{P\&L} = 2\alpha - 1.$$

The expected shortfall $ES_{\alpha}^{P\&L}(t, T)$ is

$$-\frac{1}{1-\alpha} \int_{-1}^{1-2\alpha} x \frac{1}{2} dx = -\frac{1}{1-\alpha} \frac{x^2}{4}\bigg|_{-1}^{1-2\alpha} = -\frac{1}{1-\alpha}\left(\frac{(1-2\alpha)^2}{4} - \frac{1}{4}\right).$$

For example if $\alpha = 0.95$, then $VaR_{0.95}^{P\&L} = 0.9$ and

$$ES^{P\&L} = \frac{1}{1-0.95}\left(\frac{(-0.9)^2}{4} - \frac{1}{4}\right) = 0.95.$$

This means that by entering into the swap we can lose no more tha 0.9 millions with 95% probability. But if we lose more than this amount, in average we lose 0.95 millions.

Example 10 (Expected Shortfall calculation for a discrete random variable)
Consider example 5. The P&L distribution conditional to have a loss larger than 0.5 is given by

P&L|P&L ≤ -0.5	-2	-1	-0.5
Cond. Prob.	$\frac{0.04}{0.04+0.05+0.2}$	$\frac{0.05}{0.04+0.05+0.2}$	$\frac{0.2}{0.04+0.05+0.2}$

The ES is given by

$$\left(\frac{4}{29} \times (-2) + \frac{5}{29} \times (-1) + \frac{2}{29} \times (-0.5)\right) = \frac{21}{29} = 0.7241.$$

Example 11 (Expected Shortfall is subaddditive) *With reference to example 6 on the defaultable corporate bonds A and B, we can compute the 95% Expected Shortfall. We obtain, using the P&L distribution in table 3 with reference to the two bonds taken separately*

$$ES_{0.95}^A = ES_{0.95}^B = -\frac{(-28.9) \times 0.03 + (-8.9) \times 0.02}{0.05} = 20.9\$.$$

With reference to the portfolio expected shortfall, we have

$$ES_{0.95}^{A+B} = 27.8\$.$$

If we use ES as risk-measure, the subadditivity holds:

$$ES_{0.95}^{AB} = 27.8\$ \leq ES_{0.95}^A + ES_{0.95}^B = 41.8\$.$$

The subadditivity is a desirable property: it creates incentives to portfolio diversification, since it appears to have less risk. The subadditivity of ES holds always, as shown in Artzner et al. [2].

Example 12 (VaR is subaddditive if returns are Gaussian) *As we will see shortly, under normality, the VaR depends on the volatility of the log-return. If the portfolio is made of two assets, we have*

$$\sigma^2\left(X+Y\right) = \sigma^2\left(X\right) + \sigma^2\left(Y\right) + 2\rho\sigma\left(X\right)\sigma\left(Y\right),$$

where ρ is the correlation coefficient between the two risky asset. Assuming, without loss of generality, zero expected return, the VaR of the individual positions and of the portfolio is

$$VaR^X = -z_{1-\alpha}\sqrt{\sigma^2\left(X\right)}, VaR^Y = -z_{1-\alpha}\sqrt{\sigma^2\left(Y\right)},$$
$$VaR^{X+Y} = -z_{1-\alpha}\sqrt{\sigma^2\left(X\right) + \sigma^2\left(Y\right) + 2\rho\sigma\left(X\right)\sigma\left(Y\right)}.$$

Subadditivity holds if

$$VaR^{X+Y} \leq VaR^X + VaR^Y$$

i.e.

$$\sqrt{\sigma^2\left(X\right) + \sigma^2\left(Y\right) + 2\rho\sigma\left(X\right)\sigma\left(Y\right)} \leq \sqrt{\sigma^2\left(X\right)} + \sqrt{\sigma^2\left(Y\right)} = \sigma\left(X\right) + \sigma\left(Y\right),$$

i.e. iff

$$\sigma^2\left(X\right) + \sigma^2\left(Y\right) + 2\rho\sigma\left(X\right)\sigma\left(Y\right) \leq \sigma^2\left(X\right) + \sigma^2\left(Y\right) + 2\sigma\left(X\right)\sigma\left(Y\right),$$

and this is true if and only if the correlation coefficient ρ between the two risky asset is ≤ 1, that is always the case.

Remark 2 The VaR does not describe the maximum loss, but the maximum loss at a given probability level. This means that there is a $1 - \alpha$ probability of having losses larger than the VaR. This is illustrated in the top panel of Figure 4, in which we notice that the 95% VaR is 15 ml USD, but the maximum loss is 30 ml USD.

Remark 3 VaR does not describe the losses in the left tail. In Figure 4 we have plotted two probability distributions having the same 95% VaR equal to 15 ml USD, but with a very different behavior of the tails. The distribution in the top panel has a probability mass of 5% concentrated at -30 and -15, whilst the distribution in the bottom panel has a left tail slowly decaying. This means that if our loss is larger than 15, this loss will be equal to 15 or 30 USD. In the second case, it can assume values 15 or 20 USD. The different behavior of the tails is captured by the ES. In the first case, if the loss is larger than 15, the expected loss equals to

$$\frac{2}{5} \times 30 + \frac{3}{5} \times 15 = 21 mlUSD.$$

In the second case

$$\frac{2}{5} \times 20 + \frac{3}{5} \times 15 = 17 mlUSD.$$

There are three elements in VaR (and ES) calculations.

First, the probability of losses exceeding VaR, $1 - \alpha$, needs to be specified, with the most common probability level being 1%. Theory provides little guidance about the choice of α. VaR levels of 1%–5% are very common in practice, but less extreme higher numbers (e.g., 10%) are often used in risk management on the trading floor, and more extreme lower numbers (e.g., 0.1%) may be used for applications like long-run risk analysis for pension funds.

Second, the holding period (i.e., the time period over which losses may occur). This is usually one day, but can be longer or shorter depending on particular circumstances. Those who actively trade their portfolios may use a one-day holding period, but longer holding periods are more realistic for institutional investors and non-financial corporations. Many proprietary trading desks focus on intraday VaR, from one hour to the next.

Third and final step is identification of the probability distribution of the profit and loss of the portfolio. This is the most difficult and important aspect of risk modeling. The standard practice is to assume a probability model generating the data and then estimate it by using past observations. The two most common approaches are the parametric and non parametric methods. Among the parametric approaches, the most well known is the
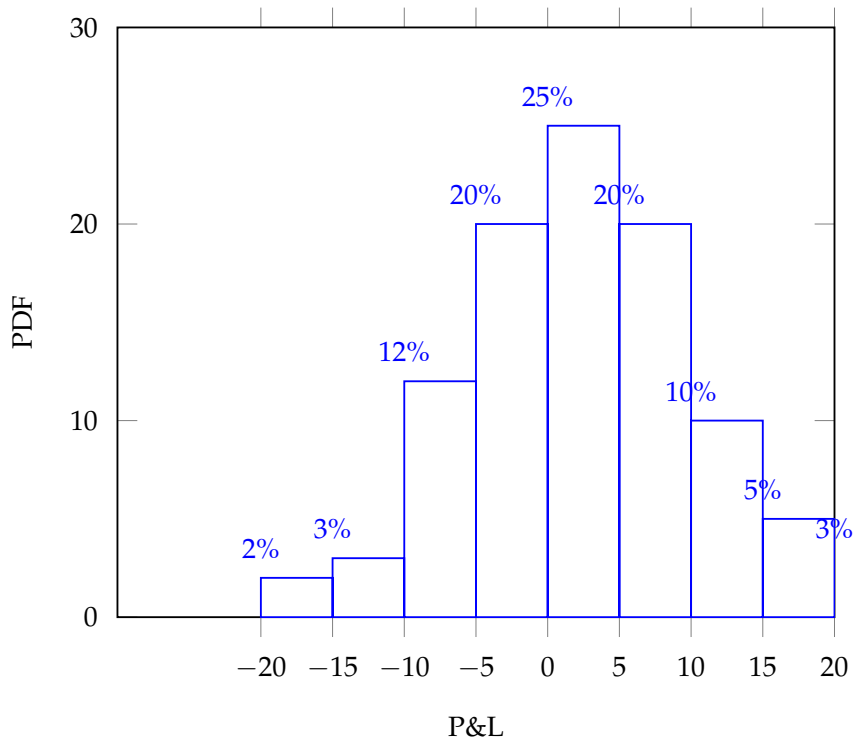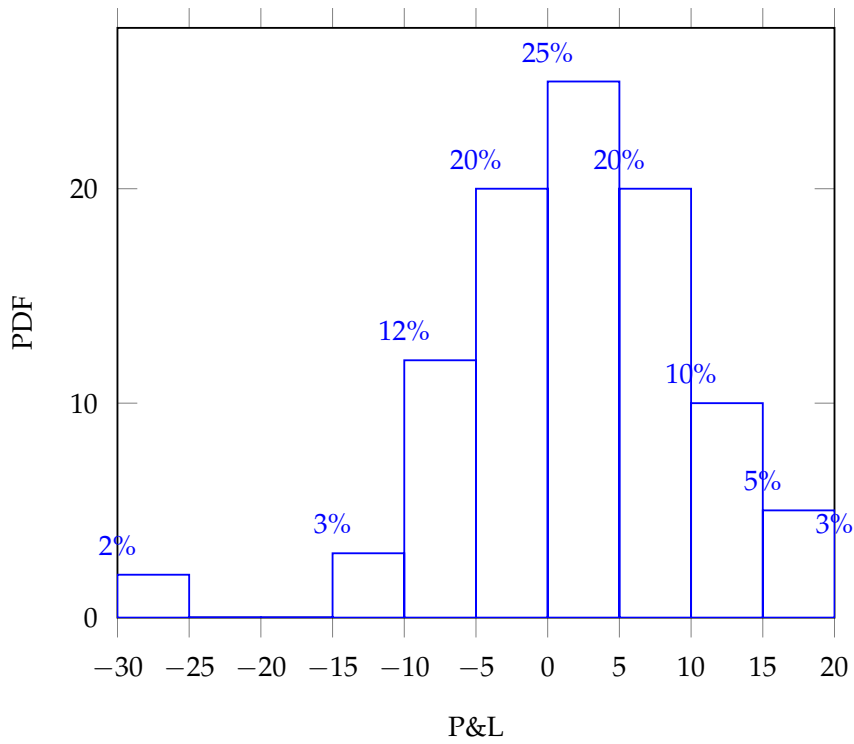
Figure 4: Top Panel: the distribution has a 95% VaR equal to 15. The maximum loss is 30. **Bottom Panel:** the distribution has a 95% VaR equal to 15. The maximum loss is 20.

15

Gaussian one. Among the non-parameteric approaches, the most commonly used is the so called historical simulation or bootstrap approach. We briefly review them in the following sections.

3   The P&L Distribution

In the VaR definition, the most important element is the profit and loss (or log-return) distribution at the time horizon $T - t$. We concentrate now the attention on the construction of the P&L distribution in general terms and then we will cover specific examples related to energy markets.

The two simplest approaches for linear positions (such as commodities) are

1. The *parametric approach*. In this case we require an assumpton regarding the distribution of log-returns. For example, we can assume that they are distributed according to a Gaussian distribution. Therefore, the estimation of the probability distribution requires the estimation of the parameters (mean and standard deviation in the Gaussian case) of the population. Cleary, other parametric distributions can also be considered.

2. The *non-parametric approach*. In this case, we do not make any model assumption and we estimate the unknown population distribution using the empirical distribution (by example using histograms).

Clearly, combinations and improvements of these two basic approaches are also possible.

3.1   Parametric VaR under the normality assumption

If we assume that log-returns have a Gaussian distribution, so that

$$r\left(t, t+\Delta\right) \sim \mathcal{N}\left(\mu_\Delta, \sigma_\Delta^2\right), \tag{4}$$

where $\mu_\Delta = \mathbb{E}\left(r(t, t+\Delta)\right)$ is the expected return of Δ-period log-return and $\sigma_\Delta^2 = \mathbb{V}ar\left(r(t, t+\Delta)\right)$ is the variance, it can be easily shown that the following result holds.

Fact 13 (Gaussian log-returns, VaR and ES) *If log-returns over a period of length Δ follow a Gaussian distribution, then the VaR with horizon Δ is*

$$VaR_\alpha\left(t, t+\Delta\right) = -\left(\mu_\Delta + z_{1-\alpha}\sigma_\Delta\right), \tag{5}$$

where z_α is the $\alpha\%$ quantile of the standard Gaussian distribution, i.e.

$$\Phi\left(z_{1-\alpha}\right) = 1 - \alpha.$$

The Expected Shortfall is given by

$$ES_\alpha^r\left(t, t + \Delta\right) = -\left(\mu_\Delta - \sigma_\Delta \frac{\phi\left(z_{1-\alpha}\right)}{1-\alpha}\right), \tag{6}$$

where $\phi\left(x\right)$ is the standard Gaussian PDF, i.e.

$$\phi\left(x\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

If Δ-period log-returns have zero-autocorrelation[5], it follows that cumulative log-returns are also Gaussian

$$r\left(t, t + n\Delta\right) \sim \mathcal{N}\left(\mu_\Delta n, \sigma_\Delta^2 n\right),$$

so that the corresponding VaR is

$$VaR_\alpha\left(t, t + n\Delta\right) = -\left(\mu_\Delta n + z_{1-\alpha}\sigma_\Delta \sqrt{n}\right). \tag{7}$$

and similarly the ES is

$$ES_\alpha^r\left(t, t + \Delta\right) = -\left(\mu_\Delta n - \sigma_\Delta \sqrt{n}\frac{\phi\left(z_{1-\alpha}\right)}{1-\alpha}\right). \tag{8}$$

An example of computation of VaR and ES under the parametric Gaussian approach for typical parameter values and for varying confidence level α is given in Table 4. Notice that the Expected Shortfall is always larger than the VaR at the same confidence level.

In the left panel of Figure 5 we plot the so called *term structure of VaR and ES*, i.e. the values of the two risk-measures at different time horizons. Notice that if $\mu > 0$, there is an horizon at which the VaR is maximum and is given[6] by $n^\star = \left(\sigma_\Delta z_{1-\alpha}/(2\mu_\Delta)\right)^2$. This is due to the fact that the expected return grows linearly with the horizon, whilst the volatility increases according to the much slower rate of the square root rule. This is shown in the right panel of Figure 5. In Figure 6 we plot the VaR and ES for different confidence levels.

---

[5]This assumption, sometimes also called zero serial correlation, means that there is no linear dipendence between past returns and current returns.

[6]The horizon at which the ES is maximized is given by the same formula, except for $z_{1-\alpha}$ being replaced by $\phi(z_{1-\alpha})$.

μ	σ	α	$z_{1-\alpha}$	VaR_α	ES_α^r
4%	20%	90%	-1.2816	22%	31%
4%	20%	93%	-1.4395	25%	34%
4%	20%	95%	-1.6449	29%	37%
4%	20%	98%	-1.96	35%	43%
4%	20%	99%	-2.3263	43%	49%
4%	50%	90%	-1.2816	60%	84%
4%	50%	93%	-1.4395	68%	90%
4%	50%	95%	-1.6449	78%	99%
4%	50%	98%	-1.96	94%	113%
4%	50%	99%	-2.3263	112%	129%

Table 4: Computing VaR and ES under the Gaussian assumption. Typical annualized values for the expected return and the volatility are used.
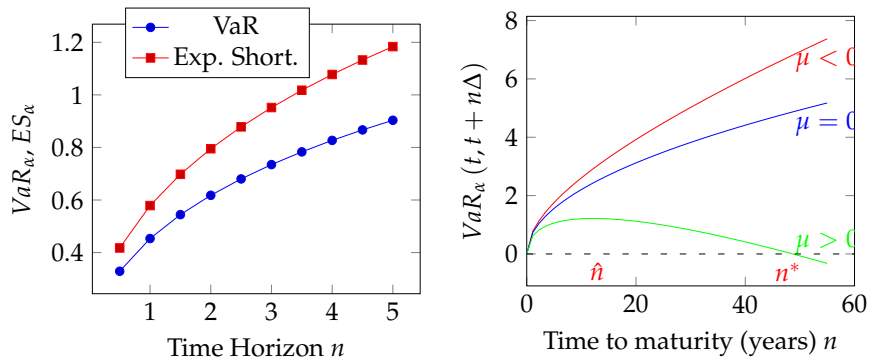
Figure 5: Left panel: Term Structures of (Gaussian) VaR and ES versus time horizon (Parameters: $\mu = 1\%, \sigma = 30\%, \alpha = 95\%$). **Right panel:** Gaussian VaR and ES versus confidence level (Parameters: $\mu = 1\%, \sigma = 30\%, n = 1$ year).
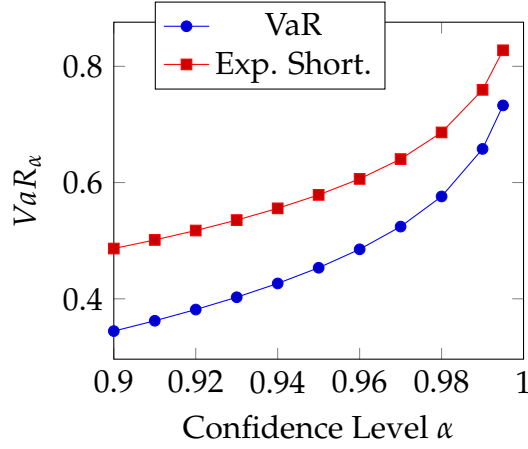
Figure 6: VaR and ES as function of the confidence level α in the parametric Gaussian model ($\mu = 0.04$ and $\sigma = 0.3$).

Estimating μ and σ   In practice, to make formula (7) and (8) operational, we need to estimate the population mean and standard deviation. The simplest approach is to use the sample mean $\hat{\mu}_\Delta$ and the sample standard deviation $\hat{\sigma}_\Delta$ of a sample of log-returns. Given a sample of historical log-returns $r_0, ..., r_{T-1}$, where $(r_j = r(j\Delta, (j+)\Delta)$, the sample mean and the sample variance[7], referred to a period of length Δ, are computed according to

$$\widehat{\mu}_\Delta = \frac{1}{T} \sum_{j=0}^{T-1} r_j, \text{ and } \widehat{\sigma}_\Delta^2 = \frac{1}{T} \sum_{j=0}^{T-1} \left(r_j - \widehat{\mu}_\Delta\right)^2.$$

Then the sample standard deviation is obtained as the square root of the sample variance. So instead of using formula (5) we use

$$\widehat{VaR}_\alpha (t, t + \Delta) = - \left(\widehat{\mu}_\Delta + z_{1-\alpha}\widehat{\sigma}_\Delta\right). \tag{9}$$

Formula (9) can be computed in Excel according to
```
=-(Mean+Volatility*NORMSINV(1-Confidence)).
```
Here `Mean` is computed using the Excel function `Average` and `Volatility` using the square root of the sample variance that is obtained via the Excel function `VAR.P`. In Excel, the ES formula can be computed via
```
=-Mean+Volatility*NORMDIST(NORMSINV(Confidence),0,1,FALSE)/(1-Confidence).
```

---

[7]In the definition of the sample variance below, we divide by T rather than by $T - 1$, because we prefer to use a biased estimator, but with a smaller mean square error. In addition, for the sample sizes used the difference is negligible.

Example 14 (Estimating VaR via the Gaussian parametric approach) *Let us consider the price series of New York Harbor Conventional Gasoline Regular Spot Price FOB (Dollars per Gallon) in August 2015 reported in Table 5 and the corresponding daily log returns series. Assuming Gaussian daily returns, we estimate*

Date	Price	Log-return	Date	Price	Log-return
Aug 03, 2015	1.751				
Aug 04, 2015	1.764	0.0074	Aug 18, 2015	1.644	-0.0151
Aug 05, 2015	1.674	-0.0524	Aug 19, 2015	1.56	-0.0524
Aug 06, 2015	1.655	-0.0114	Aug 20, 2015	1.537	-0.0149
Aug 07, 2015	1.631	-0.0146	Aug 21, 2015	1.556	0.0123
Aug 10, 2015	1.705	0.0444	Aug 24, 2015	1.485	-0.0467
Aug 11, 2015	1.713	0.0047	Aug 25, 2015	1.456	-0.0197
Aug 12, 2015	1.772	0.0339	Aug 26, 2015	1.386	-0.0493
Aug 13, 2015	1.729	-0.0246	Aug 27, 2015	1.464	0.0548
Aug 14, 2015	1.699	-0.0175	Aug 28, 2015	1.524	0.0402
Aug 17, 2015	1.669	-0.0178	Aug 31, 2015	1.651	0.0800

Table 5: Data for Example 14. Source: `www.eia.gov/dnav/pet/pet_pri_spt_s1_d.htm`

the population mean using the sample mean

$$\widehat{\mu}_\Delta = \frac{0.0074 - 0.0524 - 0.0114 + \cdots + 0.0402 + 0.0800}{20} = -0.0029,$$

and the population variance using the sample variance

$$\widehat{\sigma}_\Delta^2 = \frac{0.0074^2 + 0.0524^2 + 0.0114^2 + \cdots + 0.0402^2 + 0.0800^2}{20} - (0.0029)^2$$
$$= 0.0365^2.$$

Therefore, the estimated 95% VaR at 1 day horizon is obtained using the plug-in estimate

$$\widehat{VaR}(t, t+\Delta) = -(-0.0029 - 1.6448 \times 0.0365) = 0.0630.$$

The estimated ES is

$$\widehat{ES}(t, t+\Delta) = -\left(-0.0029 - \frac{e^{-\frac{(-1.6448)^2}{2}}}{\sqrt{2\pi}(1-\alpha)} \times 0.0365\right) = 0.0783.$$

This means that we cannot lose more than 6.30% over the next day at a confidence level of 95%. If we lose more than 6.30%, the expected loss is 7.83%. If we are interested in the 10 days VaR, we obtain

$$\widehat{VaR}(t, t+10\Delta) = -\left(-0.0029 \times 10 - 1.6448 \times 0.0365 \times \sqrt{10}\right) = 0.2194,$$

i.e. we cannot lose more than 22% over the next 10 days at a confidence level of 95%.

3.1.1 Parametric VaR and estimation risk

As illustrated in last example, the reported VaR number is a an estimate of the true VaR and therefore it is affected by the so called *estimation error*, i.e. it has a sampling variability due to the limited sample size. The issue is how large the sampling error is. Indeed, from the viewpoint of VaR users, it is useful to assess the degree of precision in the reported VaR.

Assuming that the population is Gaussian with zero mean and unknown variance σ^2, it can be shown that, for large sample sizes T, the standard error, which is a measure of the quality of an estimator, of the sample standard deviation estimator approaches

$$\lim_{T\to\infty} s.e. \left(\widehat{\sigma}_\Delta\right) = \sigma_\Delta \frac{1}{\sqrt{2T}},$$

(this is called asymptotic (i.e. for large T) standard error of the sample standard deviation). Therefore the accuracy of the VaR estimator can be assessed as

$$\lim_{T\to\infty} s.e. \left(\widehat{VaR}_\Delta\right) = \sigma_\Delta |z_{1-\alpha}| \sqrt{\frac{n}{2T}}. \tag{10}$$

This means that the accuracy of the VaR estimate deteriorates[8] for

- large absolute values of $z_{1-\alpha}$, i.e. for very high or very low values of the confidence level;

- long horizon n: it is more difficult to forecast at longer horizons than at shorter ones;

- large population volatility σ_Δ: the larger the volatility of returns, the lower the quality of the VaR estimate;

- small sample size T: we have less information to estimate accurately the unknown parameters.

The above formula also allows us to build confidence intervals around the VaR estimate. For example, a 95% confidence interval for VaR is given by

$$\widehat{VaR}_\alpha \pm 1.645 \times \widehat{\sigma}_\Delta |z_{1-\alpha}| \sqrt{\frac{n}{2T}}.$$

---

[8]Similar remarks hold for the ES.

We note that particular caution should be used when comparing VaR figures at different confidence levels α, for a given sample size T, because they are not estimated with the same accuracy. For example, let us assume that the population is Gaussian with zero mean and unknown variance σ^2 and that we have a 90% VaR estimate obtained using 250 daily observations. According to formula (10), the accuracy of the 90% VaR estimate is given by

$$\widehat{\sigma}_\Delta |z_{1-0.9}| \sqrt{\frac{1}{2 \times 250}}.$$

If we want to estimate the VaR at a different confidence level, but with the same level of accuracy, we should have a sample of size T_α such that

$$\widehat{\sigma}_\Delta |z_{1-0.9}| \sqrt{\frac{1}{2 \times 250}} = \widehat{\sigma}_\Delta |z_{1-\alpha}| \sqrt{\frac{1}{2 \times T_\alpha}},$$

or equivalently

$$T_\alpha = 250 \times \left(\frac{|z_{1-\alpha}|}{|z_{1-0.9}|} \right)^2. \tag{11}$$

For example, in Table 6 we compute the sample sizes T_α at different confidence levels so that the accuracy of the VaR estimator is the same as the one of the VaR estimator at $\alpha = 90\%$ confidence level given above. The disappointing point is that if for example we use 250 observations for estimating the 90% VaR, we should use a sample with 412 observations to estimate the 95% VaR at the same level of accuracy. This size increases to 1,454 (i.e. approximately 6 years of data) for the 99.9% confidence level. This should raise some awareness in blindly believing the accuracy of the VaR estimate at very large confidence levels. The danger here is obvious: if VaR estimates are too inaccurate and users take them seriously, they could take on much bigger risks and lose much more than they had bargained for.

α	$z_{1-\alpha}$	$|z_{1-\alpha}|$	$\frac{|z_{1-\alpha}|}{|z_{1-0.9}|}$	$\left(\frac{|z_{1-\alpha}|}{|z_{1-0.9}|} \right)^2$	T_α
0.9	-1.2816	1.2816	1.0000	1.0000	250
0.95	-1.6449	1.6449	1.2835	1.6473	412
0.99	-2.3263	2.3263	1.8153	3.2952	824
0.999	-3.0902	3.0902	2.4113	5.8145	1454

Table 6: Computation of T_α in formula (11) for different confidence levels.

The advantage of the Gaussian parametric model is that its implementation requires the estimation of only two parameters, so that the accuracy of the VaR estimate turns out to be much higher than for competing models, such as the historical simulation method. The main problem is that the Gaussian parametric approach is a model-dependent procedure, so it provides misleading indications if the model is poor. In addition, it is not able to capture "fat-tailed" and skewed distributions. In Figure 7 we compare the Gaussian density with the empirical one of Gasoline prices. A similar comparison is also given in Figure 8 for the crude oil WTI prices. In both cases, using the Jarque-Bera test[9] we can reject the Gaussian assumption of log-returns of these commodity price series.

3.1.2 Issues in volatility estimation

The estimation procedure adopted in the Gaussian approach for estimating the population volatility is questionable, because it assigns the same weight to old and recent observations. We say that the sample variance represents an unconditional estimator of the population volatility. In simple terms, the volatility estimate does not depend on the ordering with which we have registered historical returns: reshuffling the sample does not change the variance estimate. Recent information receives the same weights as older information, although intuition suggests that recent data should be more relevant. Therefore, the **term structure of volatility forecasts**, i.e. the plot of the future 1-period expected volatility against the time horizon, is flat. Instead, a non flat volatility term structure can be related to seasonality effects in volatility. For example, Figure 9 illustrates the annualized volatility of natural gas prices at Henry Hub for different months. The highest volatility levels tend to occur from October to February, which in general are the coldest months of the year. Moreover, fluctuations in demand for heating due to unpredictable weather conditions cannot always be met with storage volumes, because the natural gas in storage has to

---

[9]The Jarque and Bera [21] test measures departure from normality in terms of the skewness and kurtosis. The Jarque-Bera JB statistic is defined as

$$\widehat{JB} = \frac{T}{6} \times \widehat{sk}_\Delta^2 + \frac{T}{24} \times \left(\widehat{k}_\Delta - 3\right)^2,$$

where, $\widehat{sk}$ and $\widehat{k}$ refer to the sample estimators of the skewness and kurtosis. Under the null hypothesis of normally distributed errors, $\widehat{JB}$ has an asymptotic chi-square distribution with 2 degrees of freedom. Large values of the Jarque-Bera statistic compared to the critical level of the chi-square distribution with 2 degrees of freedom indicate departures from normality.
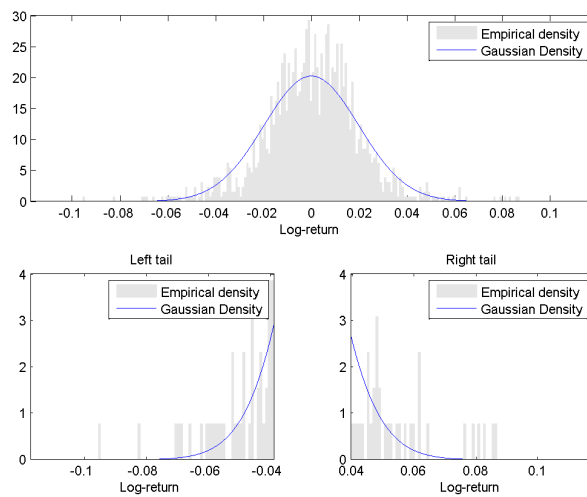
Figure 7: Top panel: Histogram of daily returns on the NY Gasoline prices in the period Jan 4th, 2010 to Aug 31st 2015 (number of observation: 1427), compared to the normal distribution with same mean and standard deviation as the historical one. **Lower left panel**: Left tail of histogram and Gaussian PDF. **Lower right panel**: Right tail of histogram and Gaussian PDF. Skewness coefficient in the data is equal to -0.65 whilst the kurtosis is equal to 5.16. The Jarque-Bera test rejects the null hypothesis that the sample data is originated by a Gaussian distribution.
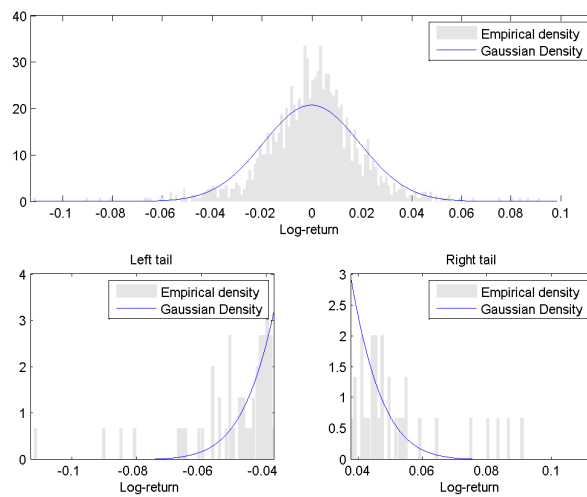
Figure 8: Top panel: Histogram of daily returns on the Cushing WTI Spot Price FOB (Dollars per Barrel) prices in the period Jan 4th, 2010 to Aug 31st 2015 (number of observation: 1427), compared to the normal distribution with same mean and standard deviation as the historical one. **Lower left panel**: Left tail of histogram and Gaussian PDF. **Lower right panel**: Right tail of histogram and Gaussian PDF. Skewness coefficient in the data is equal to -0.04 whilst the excess of kurtosis is equal to 6.70. The Jarque-Bera test rejects the null hypothesis that the sample data is originated by a Gaussian distribution.
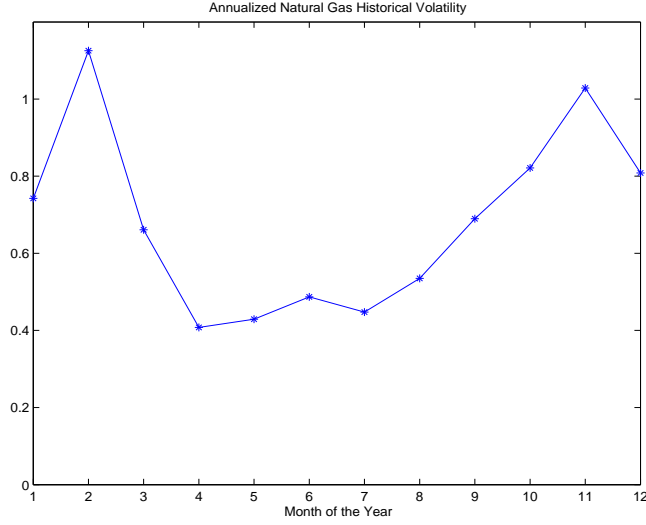
Figure 9: Annualized historical volatility by month of Henry Hub Natural Gas Prices. Observation period: January 2007 to September 2015.

serve the entire winter season. Hence, in this period we also register more erratic price movements.

In the case in which, we expect a predictable movement in volatility, the model in formula (4) has to be replaced with

$$r(t, t + \Delta) \sim \mathcal{N}\left(\mu_\Delta, \sigma_\Delta^2(t)\right), \tag{12}$$

and therefore, again assuming zero serial correlation in returns,

$$r(t, t + n\Delta) \sim \mathcal{N}\left(\mu_\Delta, \sum_{j=0}^{n-1} \sigma_\Delta^2(t + j\Delta)\right). \tag{13}$$

Therefore in the VaR (and ES) formula, instead of using the square root rule to extrapolate the 1-period volatility to an *n*-period horizon, we compute the so called integrated variance over the period and then we take its square root. For example, if we expect the daily volatility of Natural Gas in January to be 4.70% and in February 7.12%, and on December 31st we require a VaR horizon of 40 days the integrated variance is computed according to

$$31 \times (4.70\%)^2 + 9 \times (7.12\%)^2 = 0.1140,$$

Then the VaR formula is applied using the square root of 0.1140, i.e. 33.76%.

Another possibility is to use in the VaR formula the implied volatility forward curve, i.e. the market's volatility estimate extracted from quoted
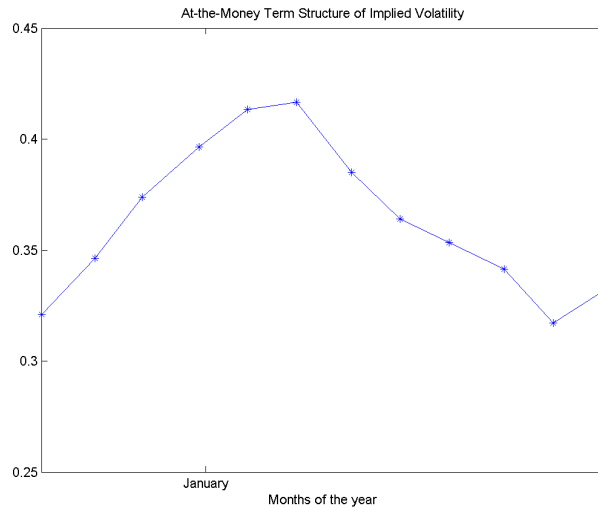
Figure 10: Term Structure of at-the-money implied volatility of Natural Gas Options (European) quoted at CME. Trade Date: Friday 18th, 2015.

prices of call and put options. Whilst the historical volatility is computed from historical prices of the underlying commodity and therefore it yields information about how prices have moved in the (recent) past, the implied volatility, eventually interpolated to the VaR horizon, provides a market expectation of future volatility. Figure 10 illustrates the term structure of at-the-money implied volatility extracted from quoted prices of natural gas options, as of September 2015. Its seasonal shape is evident. We recall that, assuming that log-returns evolve according to an arithmetic Brownian motion, the square of the implied volatility is an estimate of the integrated variance of the log-returns from the trade date up to the option expiry. Therefore, implied volatility provides a direct input to be used in the VaR formula. Unfortunately, not all commodity markets have reliable options data.

Another possibility is to use an exponential-weighted moving average (EWMA) scheme, or some kind of conditional volatility model. In this approach, the volatility at time t is dependent on the recent return history. This approach is able to capture the so called clustering (or persistence) in volatility, i.e. high (low) volatility periods are followed by high (low) volatility periods, a typical feature of financial returns as shown by the large autocorrelation of squared returns. An unconditional volatility estimator like the sample variance cannot capture this feature. In the EWMA model, the Gaussian assumption of Δ period return in (12) is augmented by a so called variance equation that describes how the next period vari-

27
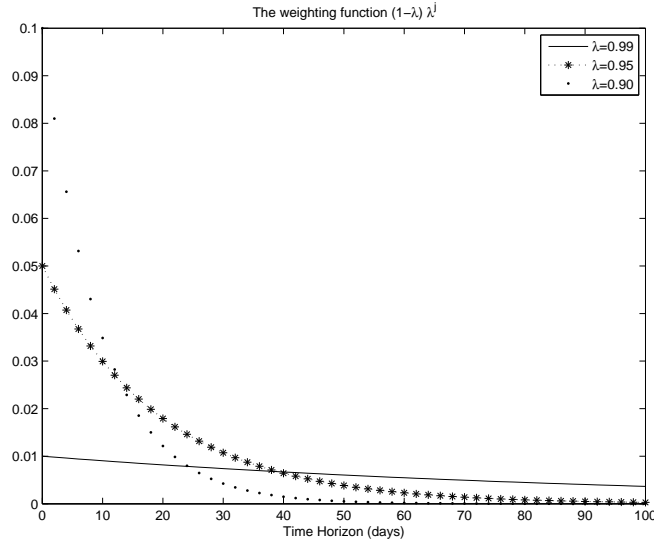
Figure 11: The Figure illustrates the behavior of the weighting function $(1 - \lambda) \times \lambda^j$ for different values of λ. Lower values of λ imply a fast decay of this function, so that only recent observations contribute to the formation of the variance forecast.

ance is related to the recent past.

In particular, given a smoothing parameter λ, $0 < \lambda < 1$, the EWMA model assumes that the variance forecast for the next period is time varying according to the rule

$$\sigma_\Delta^2(t) = (1 - \lambda) \sum_{j=0}^{\infty} \lambda^j r^2 \left(t - (j + 1) \Delta, t - j\Delta \right),$$

i.e. it is an exponentially weighted average of past squared returns. Therefore, the EWMA approach has the appealing property that the influence of any observation declines over time at a stable rate, according to the rule $(1 - \lambda) \times \lambda^j$. In practice, as we lower the value of λ, we give more weight to the most recent observations. Viceversa, as we increase the value of λ we assigns equal weight to all observations. The sample variance estimator is recovered in the limiting case of $\lambda = 1$. This is illustrated in Figure 11, where we plot the weighting function for different values of λ.

It is useful to rewrite the EWMA variance equation according to the following updating rule

$$\underbrace{\sigma_\Delta^2(t)}_{\text{today forecat}} = (1 - \lambda) \times \underbrace{r^2(t - \Delta, t)}_{\text{yesterday squared market shock}} + \lambda \times \underbrace{\sigma_\Delta^2(t - \Delta)}_{\text{yesterday forecast}}.$$

$$(14)$$

This means that

- the *higher λ, the higher the persistence in the variance* through the term $\lambda \sigma^2 (t - \Delta)$;

- the *lower λ, the higher the reaction to market shocks* through the term $(1 - \lambda) r^2 (t - \Delta, t)$.

Example 15 (Updating Variance estimate using EWMA recursion) *Let us consider the return series in the second row of Table 7. Assuming $\lambda = 0.9$, in the third row we implement the recursive updating of the daily variance forecast, starting from an hypothetical initial value of 3. In the fourth row, we have the daily volatility forecast. Computations proceed as follows*

Days k	0	1	2	3	4	5	6	7	8	9	10
$r_{k\Delta}$	2	5	5	-1	5	-5	5	-5	3	-4	-2
$\sigma^2_{(k+1)\Delta}$	3	3.1	5.29	7.26	6.63	8.47	10.12	11.612	12.95	12.56	12.9
$\sigma_{(k+1)\Delta}$	1.73	1.76	2.3	2.69	2.58	2.91	3.18	3.41	3.6	3.54	3.59

Table 7: The updating procedure of the variance in the EWMA scheme. Numbers are purely illustrative to simplify the recursive computation.

- *Given an initial value for $\sigma^2 (0) = 3$, and given the return $r(0, \Delta) = 2$, the variance forecast for the following day is computed via*

$$\sigma^2_\Delta (\Delta) = 0.9 \times 3 + 0.1 \times 2^2 = 3.1$$

and the volatility forecast is $\sqrt{3.1} = 1.7607$.

- *Then 3.1 together with the squared return on the second day provides a variance forecast at the end of the second day*

$$\sigma^2_\Delta (2\Delta) = 0.9 \times 3.1 + 0.1 \times 5^2 = 5.29.$$

The volatility forecast is $\sqrt{5.29} = 2.3$.

- *If we proceed in a similar way up to the last day of the sample, the variance forecast on the last day of our sample is obtained via*

$$\sigma^2_\Delta (10\Delta) = 0.9 \times 12 - 6 + 0.1 \times (-2)^2 = 12.9.$$

- *If on day 10, we are interested in the 1-day VaR, we can use the square root of 12.9 to forecast the VaR_α measure according to the formula*

$$VaR_\alpha (10\Delta, 11\Delta) = -z_{1-\alpha} \times \sqrt{12.9}.$$

As shown in the last example, in order to use the EWMA procedure we need

- the starting value of the recursion, i.e. $\sigma_\Delta^2(0)$; in the example, this value was set arbitrarily at 3.

- an estimate of the parameter λ; in the example, this value was set arbitrarily at 0.9.

Common practice is to set $\lambda = 0.94$ if we are dealing with daily returns, and estimate the starting value of the recursion $\sigma^2(0)$ by using the sample variance relative to additional K observations occurring before the day in which we start the recursion.

Another possibility it to estimate both parameters by maximum likelihood. Let us suppose that we have collected T past daily return starting from time 0

$$sample = \{r_0, r_1, ..., r_{T-1}\},$$

where we write r_j to denote $r(j\Delta, (j+1)\Delta)$. The sample likelihood, i.e. the probability of observing the sample, is

$$\mathcal{L}\left(r_0, r_1, ..., r_{T-1} \,\middle|\, \lambda, \sigma_\Delta^2(0)\right) = \prod_{j=0}^{T-1} \frac{1}{\sqrt{2\pi\sigma_\Delta^2(j\Delta)}} e^{-\frac{1}{2}\left(\frac{r_j}{\sigma_\Delta(j\Delta)}\right)^2},$$

where $\sigma_\Delta^2(j\Delta) = \mathbb{V}(r(j\Delta, (j+1)\Delta)))$ and it is calculated recursively according to the recursion in (14). It is convenient to consider the log-likelihood

$$\log\mathcal{L}\left(r_0, r_1, ..., r_{T-1} \,\middle|\, \lambda, \sigma_\Delta^2(0)\right) = \sum_{j=0}^{T-1}\left(-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\left(\sigma_\Delta^2(j\Delta)\right) - \frac{1}{2}\left(\frac{r_j}{\sigma_\Delta(j\Delta)}\right)^2\right).$$

Now we can maximize this expression with respect to the unknown parameters λ and $\sigma_\Delta^2(0)$. This can be done easily in Excel using the Excel Solver.

Example 16 (Computing the log-likelihood) *Let us consider the returns of Example 15. The daily likelihood relative to the first observation is*

$$\frac{e^{-\frac{(2)^2}{2\times 3}}}{\sqrt{2\pi3}} = 0.11826.$$

The daily likelihood relative to the second observation is

$$\frac{e^{-\frac{(5)^2}{2\times 3.1}}}{\sqrt{2\pi3.1}} = 0.00402.$$

30

	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	λ	0.9											
2	Day	0	1	2	3	4	5	6	7	8	9	10	
3	r	2	5	5	-1	5	-5	5	-5	3	-4	-2	
4	σ_Δ^2	3	3.1	5.29	7.261	6.635	8.471	10.124	11.612	12.951	12.556	12.900	
5	σ_Δ	1.732	1.761	2.300	2.695	2.576	2.911	3.182	3.408	3.599	3.543	3.592	
6	daily lik.	0.1183	0.0040	0.0163	0.1382	0.0235	0.0313	0.0365	0.0399	0.0783	0.0595	0.0951	Sample Log Lik
7	daily log-lik	-2.1349	-5.5169	-4.1148	-1.9791	-3.7491	-3.4628	-3.3111	-3.2214	-2.5470	-2.8212	-2.3526	-35.2109
8			=F1*F4+(1-F1)*F3^2				=NORM.DIST(K3;0;K5;FALSE)					=LN(P6)	=SUM(F6:O6)

Figure 12: Computing the sample likelihood of the EWMA model in Excel.
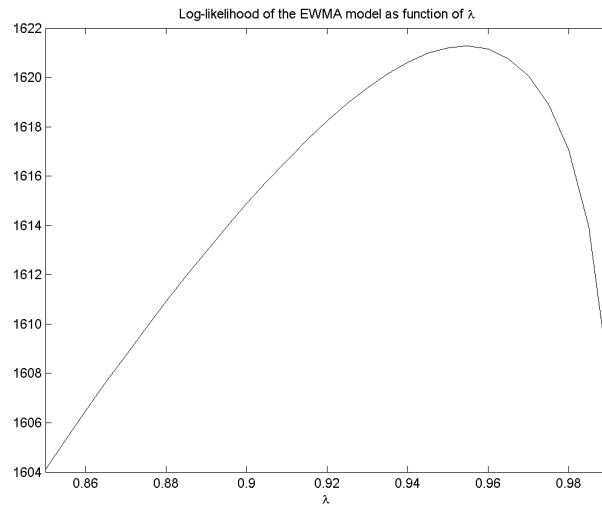
Figure 13: Log-likelihood of the EWMA model with reference to Gasoline prices for the period January 4th, 2010 to August 31st 2015.

If we consider the full sample of 10 observations, the sample log-likelihood given the starting values for the unknown parametes is obtained as

$$ln(0.11826) + ln(0.00402) + \cdots + ln(0.0951) = -35.2109.$$

In practice, we consider a much longer sample and we optimize over the values of λ and $\sigma_\Delta^2(0)$. Notice that the daily likelihood can be computed in Excel using the built-in function `NORM.DIST(return;0;volatility;FALSE)`, *where* `return` *is the observed return and* `volatility` *is the square root of the EWMA estimate of the variance. This is shown in Figure 12.*

Figure 13 illustrates how the log-likelihood changes for different values of λ. The point at which it achieves its largest value provides the maximum likelihood estimate of λ.

In the EWMA model, returns are serially uncorrelated (but not independent). Therefore the variance of the cumulative $n-$periods returns still

Time	Variance Updating	Simulated Return
t	$\widehat{\sigma}_\Delta^2(t)$ (known)	$r_{\Delta,0} = \sqrt{\widehat{\sigma}_\Delta^2(t)} \times \varepsilon_1$
$t+\Delta$	$\widehat{\sigma}_{\Delta,1}^2 = \hat{\lambda}\widehat{\sigma}_{\Delta,0} + \left(1-\hat{\lambda}\right) r_{\Delta,0}^2$	$r_{\Delta,1} = \sqrt{\widehat{\sigma}_{\Delta,1}^2} \times \varepsilon_2$
$t+2\Delta$	$\widehat{\sigma}_{\Delta,2}^2 = \hat{\lambda}\widehat{\sigma}_{\Delta,1}^2 + \left(1-\hat{\lambda}\right) r_{\Delta,1}^2$	$r_{\Delta,2} = \sqrt{\widehat{\sigma}_{\Delta,2}^2} \times \varepsilon_3$
	...	...
$t+i\Delta$	$\widehat{\sigma}_{\Delta,i}^2 = \hat{\lambda}\widehat{\sigma}_{\Delta,i-1}^2 + \left(1-\hat{\lambda}\right) r_{\Delta,i-1}^2$	$r_{\Delta,i} = \sqrt{\widehat{\sigma}_{\Delta,i}^2} \times \varepsilon_{i+1}$
	...	...
$n-1$	$\widehat{\sigma}_{\Delta,n-1}^2 = \hat{\lambda}\widehat{\sigma}_{\Delta,n-2}^2 + \left(1-\hat{\lambda}\right) r_{t+(n-2)\Delta}^2$	$r_{\Delta,n-1} = \sqrt{\widehat{\sigma}_{\Delta,n-1}^2} \times \varepsilon_{n-1}$

Table 8: Implementing Monte Carlo simulation to simulate the one period variance and the one-period log-return in the EWMA model. We use $\widehat{\sigma}_{\Delta,i}^2$ to denote $\widehat{\sigma}_\Delta^2(t+i\Delta)$, $r_{\Delta,i}$ to denote $r(t+i\Delta, t+(i+1)\Delta)$ and ε_i is a sequence of iid standard Gaussian random variables. The simulated 2-period cumulative returns is $r_{\Delta,0} + r_{\Delta,1}$. The simulated n-period cumulative returns is $r(t, t+n\Delta) = \sum_{i=0}^{n-1} r_{\Delta,i}$.

follows the square-root rule,

$$\sigma^2\left(t, t+n\Delta\right) = n \times \sigma_\Delta^2\left(t\right).$$

where $\sigma_\Delta^2\left(t\right)$ is the variance forecast of the last sample period. With reference to Example 15, in order to obtain a forecast of the 5 days variance, we compute 12.9×5.

However, if we have to compute a *n*-period VaR, we cannot use formula (7). Indeed, in the EWMA model only 1-period returns are Gaussian. Moreover, they are uncorrelated, but not independent, because they are characterized by a serial dependence through the variance dynamics. This implies that the cumulative *n*-periods log-return is no longer Gaussian. The reason is that the conditional volatility EWMA scheme generates stochastic volatility in the returns series: given different return paths, we have different volatility paths. The main effect of stochastic volatility is to generate fat tails in the distribution of cumulative returns, even if 1-period returns are Gaussian. Unfortunately, the distribution of cumulative returns is not known, and therefore the VaR formula (7) does not apply to the EWMA model. In this case, VaR estimation can be obtained only by Monte Carlo simulation. We proceed as follow. Let us suppose that we have estimated the two unknown parameters and let us call $\hat{\lambda}$ the estimated value for λ and $\widehat{\sigma}_\Delta^2\left(t\right)$ the variance forecast for the next period. Table 8 illustrates how to simulate returns given these estimates.

In practice, at each time step we have to draw, independently from the previous extractions, a new random number $\varepsilon(t + i\Delta)$ according to a standard normal random variable. We use this random draw and the variance computed according to the EWMA scheme to simulate 1-period log-return. The simulated cumulative return is obtained by the sum of the simulated one-period log-returns, given in the last column of Table 8.

Example 17 (Monte Carlo simulation of the EWMA model) *Let the estimated value of λ be 0.94. The variance estimate for the period $(t - \Delta, t)$ is 0.000585 and the corresponding log-return is -0.0087. The variance of the return over the following period is*

$$0.94 \times 0.000585 + (1 - 0.94) \times (-0.0087)^2 = 0.000555.$$

In order to simulate the log-return over $(t, t + \Delta)$, we draw a pseudo-random number ε from a standard Gaussian r.v.. Let us suppose that $\varepsilon(t) = 0.2272$. The simulated return for the period $(t, t + \Delta)$ is

$$\sqrt{0.000555} \times 0.2272 = 0.0054.$$

The variance for the second period is updated according to

$$0.94 \times 0.000555 + (1 - 0.94) \times (0.0054)^2 = 0.000523.$$

If the simulated Gaussian draw for the second period is -0.6542, the simulated one-period return for the time window $(t + \Delta, t + 2\Delta)$ is

$$\sqrt{0.000523} \times (-0.6542) = -0.0150.$$

The simulated two-days cumulative return is $0.0054 - 0.0150 = -0.0096$. Additional steps are given in Table 9.

Fact 18 (Monte Carlo Simulation and VaR in the EWMA model) *The correct approach for VaR estimation in the EWMA model is as follows*

1. *Run M simulations to obtain M simulated values of $r(t, t + n\Delta)$.*

2. *Estimate the VaR using the empirical percentile of the simulated log-returns at the desired confidence level $1 - \alpha$.*

Simulated paths of the EWMA model are illustrated in Figure 14. A comparison between the simulated distribution and the Gaussian one with same variance is given in Figure 15.

33

EWMA Monte Carlo Simulation				
Day i	$\epsilon(t)$	Variance	Log-return	Cumulative return
0		0.000585	-0.0087	
1	0.2272	0.000555	0.0054	0.0054
2	-0.6542	0.000523	-0.0150	-0.0096
3	-1.0886	0.000505	-0.0245	-0.0341
4	0.7960	0.000511	0.0180	-0.0161
5	-1.1254	0.000500	-0.0252	-0.0412

Table 9: A simulated path of daily and cumulative log-returns.

Simulated MC trajectories for 1-year cumulative return in the EWMA m
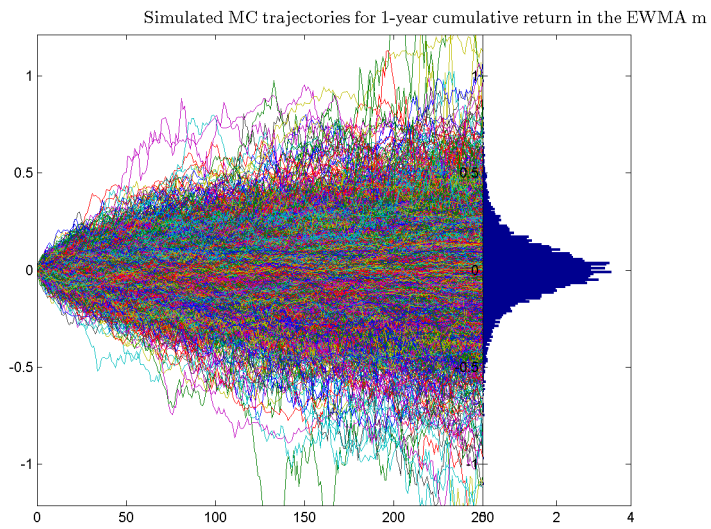
Figure 14: Simulated Paths of the EWMA Model.

34
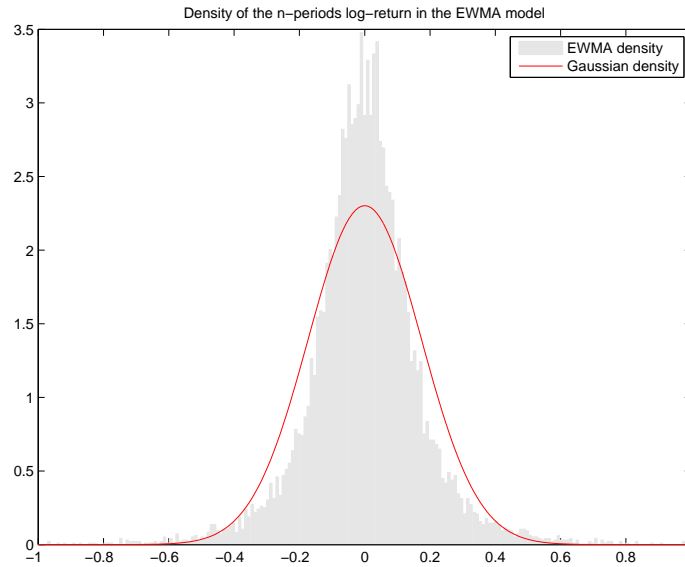
Figure 15: Simulated distribution of the EWMA Model and superimposed Gaussian distribution with same variance.

Due to the fat tails generated by the stochastic volatility effect, if we compare VaR computed using the square-root formula and Monte Carlo simulation, we can verify that for large values of the confidence level, the EWMA VaR estimate is larger than in the Gaussian case. For low values of α, the VaR computed by Monte Carlo simulation of the EWMA model is lower than in the Gaussian case

In conclusion, the EWMA is a parsimonious model able to capture the relative importance of recent returns in the volatility estimate. In addition, it also generates fat tails distribution through the conditional stochastic volatility effect. Moreover, its estimation via the maximum likelihood procedure is straigthforward. On the other side, given that innovations are Gaussian, the model is not able to generate skewed distributions. Further, it provides counterfactual longer-horizon forecasts: shocks will persist for a long time. Possible improvements to the EWMA scheme, allowing for non Gaussian innovations and generating skewed distributions via the negative relationship between variance and returns, are provided by the class of Generalized autoregressive conditional heteroskedasticity (GARCH) models. Applications of GARCH models to oil, natural gas, and electricity price are discussed in Efimova and Serletis [18].

3.2 Non-parametric VaR: Historical Simulation

Historical Simulation provides an important alternative to the parametric approach presented in the previous section, having as main advantage that no assumption is made on the model generating returns. Given a sample of size T of daily returns, $(r_1, ..., r_T)$, the non-parametric approach consists in estimating the probability density function by the histogram (empirical distribution) of observed returns. The empirical cumulative distribution function is a step function that moves up by steps of size $1/T$, T being the sample size, at each one of the T data points. If some of the r_i coincide, then that common value will receive the appropriate multiple probability mass. If $(r_1, ..., r_T)$ is the sample of iid returns, the empirical CDF is defined as follows

$$\widehat{F}_T(r) = \frac{1}{T} \sum_{i=1}^{T} 1_{r_i \leq r},$$

where $1_{\mathcal{A}}$ is the indicator function, ie it returns 1 if the event $\mathcal{A}$ is true, 0 otherwise. The VaR is then estimated by looking for the point at which the empirical CDF crosses the $(1 - \alpha)$-percentile of this empirical distribution. Given that the empirical CDF is a step function, sometimes an interpolation between adjacent points is necessary (different softwares employs different interpolation). In practice, the VaR is estimated by the empirical quantile, ie by sorting the observed returns in increasing order and selecting the one in position $(1 - \alpha)T$.

Fact 19 (Computing the VaR using historical simulation) *Let $(r_1, ..., r_T)$ be a sample of independent and identically distributed Δ-period returns. Consider the ordered returns*

$$r_{(1)} \leq r_{(2)} \leq ...r_{(T-1)} \leq r_{(T)},$$

with $r_{(1)}$ being the sample minimum and $r_{(T)}$ the sample maximum. The empirical VaR estimator at the desired confidence level is the ordered return in position $(1 - \alpha)T$

$$-\widehat{VaR}_\alpha(t, t + \Delta) = r_{((1-\alpha)T)}.$$

If $(1 - \alpha)T$ is not an integer number, we use linear interpolation with weights $1 - \gamma$ and γ

$$-\widehat{VaR}_\alpha(t, t + \Delta) = (1 - \gamma) \times r_{(\lfloor (1-\alpha)T \rfloor)} + \gamma \times r_{(\lfloor (1-\alpha)T \rfloor + 1)},$$

where $\lfloor \alpha T \rfloor$ is the integer part of αT and $\gamma = (1 - \alpha)T - \lfloor (1 - \alpha)T \rfloor$.

Position	Sorted Return	Position	Sorted Return
1	-5.245%	11	-1.461%
2	-5.237%	12	-1.141%
3	-4.927%	13	0.468%
4	-4.670%	14	0.740%
5	-2.457%	15	1.229%
6	-1.972%	16	3.386%
7	-1.782%	17	4.017%
8	-1.750%	18	4.437%
9	-1.509%	19	5.475%
10	-1.485%	20	8.004%

Table 10: Sorted log-returns of New York Harbor Conventional Gasoline Regular Spot Price FOB (Dollars per Gallon) in August 2015.

Example 20 (VaR via Empirical Quantile) *Let us consider the same sample as in Example 14. The ordered returns of Gasoline are given in Table 10. If we are interested in the 90% VaR, we have to consider the ordered return in position $(1 - 0.9) \times 20 = 2$, so that the empirical VaR estimator is the second smallest return, ie*

$$\widehat{VaR}_{0.90}\left(t, t + \Delta\right) = 5.237\%.$$

If we are interested in the 92.5% VaR level, we observe that $(1 - 0.925) \times 20 = 1.55$ is not an integer number. Hence, we use linear interpolation of the ordered return at position 1 and 2 with weights $1 - \gamma$ and γ, with

$$\gamma = (1 - 0.925) \times 20 - \lfloor (1 - 0.925) \times 20 \rfloor = 1.5 - 1 = 0.5,$$

so that

$$\widehat{VaR}_{0.925}\left(t, t + \Delta\right) = (1 - 0.5) \times 5.245\% + 0.5 \times 5.237\% = 5.241\%.$$

If the confidence level is larger than 95%, we set the negative VaR equal to the smallest registered sample value.

Unfortunately, in this case there is no simple procedure to extrapolate a n-period VaR from the 1-period one. Indeed, the square root rule is valid only for extrapolating the standard deviation under the zero autocorrelation assumption, and it is a not valid rule to project the 1-period quantile to n-periods. A possibility is to perform the above estimation looking at n-period returns, but this implies a loss of data. For example, in a sample of 250 daily returns we have only 12 monthy returns, so that the estimation

of the monthly VaR is very inaccurate. A different possibility is to estimate the VaR via Bootstrap, also known as Historical Simulation.

The bootstrap technique draws a sample of the same size as the original data set and records the VaR from the simulated sample. This procedure is repeated over and over to obtain multiple sample VaRs. In practice, this procedure is like sampling with replacement. The best VaR estimate from the full data set is the average of all sample VaRs.

If we are interested in the VaR for a n-period horizon, the procedure works as follows

1. Create a $T \times 1$ database of historical returns r on a stock.

2. Randomly select a date $t \in [1, T]$ and use the return for that date to generate a possible future return.

3. Repeat n times and obtain a simulated cumulative return as sum of the n values of the simulated 1-period returns.

4. Repeat steps 2-3 to generate T paths to obtain an estimate of the return distribution of the portfolio.

5. Compute the VaR as empirical quantile of the T simulated values.

6. Repeat steps 2-5 M times and then take the average of all sample VaRs.

This procedure relies on the independence of successive returns. However, this basic scheme can be improved in order to take into account both serial dependence and varying volatility, as well as seasonal components.

Example 21 (Bootstrap Simulation) *Let us suppose that our sample consists of 10 daily observations, see Table 11, and we are interested in simulating the cumulative return over the next 3 days.*

1. *As we have 10 daily log-returns, let us randomly draw an integer uniform in the interval [1:10],*
$$day \sim U([1, 10]),$$
   say 8. We collect the return at position 8 in Table 11, ie -0.0246.

2. *Let us make an additional extraction from the interval [1:10] and assume this is 4. Therefore, from Table 11, the simulated return on the second day is -0.0146.*

38

Day	Price	Log-ret.
Aug 03, 2015	1.751	
Aug 04, 2015	1.764	0.740%
Aug 05, 2015	1.674	-5.237%
Aug 06, 2015	1.655	-1.141%
Aug 07, 2015	1.631	-1.461%
Aug 10, 2015	1.705	4.437%
Aug 11, 2015	1.713	0.468%
Aug 12, 2015	1.772	3.386%
Aug 13, 2015	1.729	-2.457%
Aug 14, 2015	1.699	-1.75%
Aug 17, 2015	1.669	-1.78%

Table 11: Sample of daily log-returns used to perform historical simulation in Example 21.

3. *If we make a third extraction and we obtain 7, the simulated return on the third day is 0.0339.*

4. *Therefore, the simulated 3-days return is*

$$-0.0246 - 0.0146 + 0.0339 = -0.0053.$$

5. *Let us repeat steps 1-4 as many times as the number observations in our sample, ie 10 times. Given that we need to simulate 3 returns in order to generate each return paths, in total we have to extract $3 \times 10 = 30$ integer random numbers.*

6. *We end up with a simulated series of returns, as reported in Table 12. The first column refers to the number of simulations that has to be equal to the sample size T. Columns 2-4 refer to the simulation of 30 integer numbers in the range [1:10]. Columns 5-8 refer to the boostrapped daily returns. In particular, column 5 refers to simulated returns for the first day. Column 6 refers to simulated returns for the second day. Column 7 refers to simulated returns for the third day. The last column refers to the simulated three day returns that is obtained by summing the three 1-day returns. If we are interested in the 90% VaR, we consider the smallest return among the T simulated 3-day returns. It is given by -9.15%. Therefore the 3 days VaR is 9.15%.*

In general, this procedure is repeated a few times, a suggested typical number falls between 500 and 1000. Each simulation returns an estimated VaR

Sim.	1	2	3	$r(t, t+\Delta)$	$r(t+\Delta, t+2\Delta)$	$r(t+2\Delta, t+3\Delta)$	$r(t, t+3\Delta)$
1	9	9	8	-1.75%	-1.75%	-2.46%	-5.96%
2	7	1	5	3.39%	0.74%	4.44%	8.56%
3	8	2	4	-2.46%	-5.24%	-1.46%	-9.15%
4	2	5	4	-5.24%	4.44%	-1.46%	-2.26%
5	6	4	9	0.47%	-1.46%	-1.75%	-2.74%
6	3	3	7	-1.14%	-1.14%	3.39%	1.10%
7	10	9	1	-1.78%	-1.75%	0.74%	-2.79%
8	3	6	3	-1.14%	0.47%	-1.14%	-1.81%
9	10	4	3	-1.78%	-1.46%	-1.14%	-4.38%
10	4	3	10	-1.46%	-1.14%	-1.78%	-4.38%

Table 12: Implementing historical simulation to generate a *n*-periods simulated distribution.

number and the bootstrap estimate is the average of the simulated VaRs. For example, in Table 13 we report 500 simulated VaR. The average of the simulated VaR is 8.32% which represents the bootstrap estimate of the 3 days VaR.

Simulation	1	2	...	499	500
VaR	3.26%	7.84%	...	5.67%	8.16%

Table 13: Simulated 3 days VaR out of 500 historical simulations.

We can also use the 500 simulated VaR to build a confidence interval around the point estimate. For example, if we want to consider a 99% confidence interval for $VaR_{90\%}$, we sort the 500 simulated values, and consider the one at position $0.01 \times 500 = 5$ and the one at position $0.99 \times 500 = 495$. In our simulation, those values are 12.93% and 3.26%. In conclusion, the 99% confidence interval for the bootstrap estimate is (3.26%, 12.93%). The length of this interval appears quite large due to the limited sample size, just 10 observations.

In general, for a given sample size, the confidence interval length is much larger than in the parametric Gaussian case. Indeed, it can be shown that the asymptotic standard error of the sample quantile is given by

$$s.e. \left(V\hat{a}R_{1-\alpha} \right) = \sqrt{\frac{\alpha(1-\alpha)}{T\hat{f}_r^2(-V\hat{a}R_{1-\alpha})}},$$
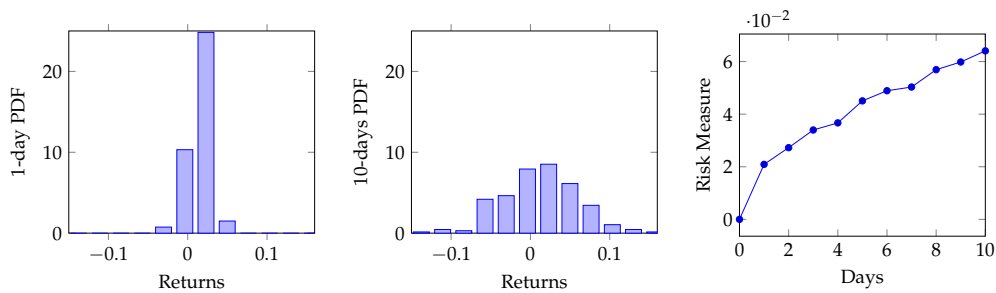
Figure 16: Left panel: Simulated 1-day log-return distribution. **Middle panel**: Simulated 10-days log-return distribution. **Right panel**: VaR term structure.
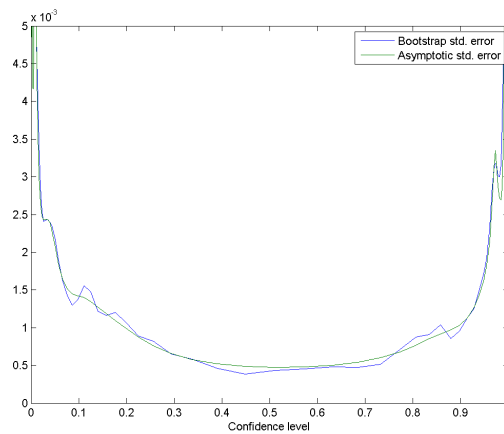
Figure 17: Asymptotic standard error of bootstrap estimator of the quantile: analytical formula vs Monte Carlo based estimate.

where $\hat{f}_r$ is the (estimated) density of log-return. This result, stated for example in Jorion's book, can be verified by using the bootstrap simulation approach as follows

1. Resample M times a sample of size T from the original sample;

2. For each sample compute the VaR at the desired confidence level;

3. Compute the standard deviation of the M estimates.

This standard deviation provides the accuracy of the bootstrap estimate and if the sample size is large enough it should agree with asymptotic standard error given above, see figure 17. If T is small, the above procedure provides a valid alternative to the asymptotic estimate.

The above formula also provides an useful tool to compare to the accuracy of the parametric Gaussian VaR. Figure (18) compares the accuracy

41
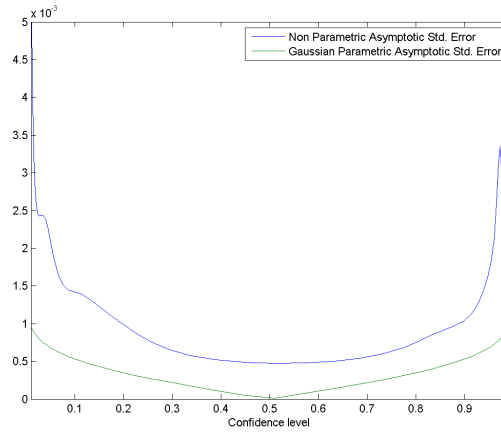
Figure 18: Comparion of the asymptotic standard error of bootstrap and parametric estimators of the quantile.

of the non parametric estimate with the Gaussian parametric one.

  a We can see that for a given confidence level and a given sample size, the parametric Gaussian estimate suffers less from the sampling error problem.

  b The non parametric one turns to be very inaccurate at very low and very high confidence levels, in practice the most interesting cases from a risk management point of view.

α	s.e.(HS)	s.e.(G)	T_{HS}
0.9	0.002337	0.000826	4000
0.95	0.002746	0.00106	3355
0.99	0.008112	0.001499	14642
0.999	0.009676	0.001992	11797

Table 14: Standard errors of the bootstrap estimate and of the Gaussian estimator. The bootstrap simulation has been repeated 100 times, generating samples of size 500. Last column gives the sample size so that the HS estimate has the same accuracy as the VaR.

On the other side, the bootstrap approach is not exposed to the model risk affecting the Gaussian approach. It capture aspects like kurtosis and skewness of the data, that is not possible adopting the Gaussian approach.

The historical simulation procedure is very simple to implement, very easy to present and is a model-free procedure. In addition, it allows the

inclusion of "fat-tailed" outcomes, although only those registered in the past. In fact, this method assumes the future will be like the past.

Finally, this approach does not capture changes in volatility, ie it only provides an unconditional distribution. However, the model can be improved as to rescale recent returns to take into account recent more volatile periods. A possibility is the so called Time Weighted Historical Simulation: instead of using equal probability weights set to $1/T$ on each simulated return, we can use probabilities which decline as we look backward in time. This method takes into account the fact that more recent observations are more relevant than old ones. It consists of the following steps.

1. The sample of the T past returns is assigned a weight $w\,(i)$ declining exponentially as

$$w\,(i) = \eta^{i-1} \frac{(1-\eta)}{1-\eta^T}, i = 1..., T$$

with $0 < \eta < 1$. Typical values for η are in the range 0.95-0.99.

2. The observations, along with the assigned weights, are sorted in ascending order.

3. The VaR at the α confidence level is the sorted return corresponding to a cumulative weight equal to $(1-\alpha)\,\%$, interpolating if necessary.

The following numerical example aims at clarifying the procedure.

Example 22 (Weighted Historical Simulation) *We proceed as follows*

1. *Assigning weigths. Let us consider a sample of 20 observations and let us set $\eta = 0.95$, so that the weight of observation i ($i=1$ is the most recent one) is*

$$w\,(i) = 0.95^{i-1} \frac{(1-0.95)}{1-0.95^{20}}, i = 1..., 20.$$

The observations and the corresponding weigths are given in the Table 15.

2. *Computing VaR. We sort the returns in increasing order and compute the cumulate weights. This is illustrated in columns 3-6 of Table 15. If we are interested in the 90% VaR, we interpolate between -1.75% (cum. weigth 6.53%) and -0.83% (cum. weigth 10.14%) to obtain*

$$VaR_{0.90}^{whs}\,(t, t+\Delta) = -\left(\frac{-1.75}{100} \times \frac{10.14-10}{10.14-6.53} + \frac{-0.83}{100} \times \frac{10-6.53}{10.14-6.53} \right)$$
$$= 0.8656\%.$$

	Returns	Weights	Sorted Returns	Weights	Cum Weights
20	0.95%	2.94%	-2.46%	3.43%	3.43%
19	-1.75%	3.10%	-1.75%	3.10%	6.53%
18	0.44%	3.26%	-0.83%	3.61%	10.14%
17	-2.46%	3.43%	-0.51%	4.43%	14.57%
16	-0.83%	3.61%	-0.24%	4.67%	19.24%
15	1.11%	3.80%	-0.24%	6.35%	25.59%
14	0.20%	4.00%	-0.20%	5.44%	31.03%
13	0.84%	4.21%	-0.08%	6.03%	37.06%
12	-0.51%	4.43%	0.10%	7.40%	44.46%
11	-0.24%	4.67%	0.20%	4.00%	48.46%
10	1.27%	4.91%	0.44%	3.26%	51.72%
9	2.29%	5.17%	0.45%	6.68%	58.41%
8	-0.20%	5.44%	0.63%	7.03%	65.44%
7	0.76%	5.73%	0.76%	5.73%	71.17%
6	-0.08%	6.03%	0.77%	7.79%	78.96%
5	-0.24%	6.35%	0.84%	4.21%	83.17%
4	0.45%	6.68%	0.95%	2.94%	86.12%
3	0.63%	7.03%	1.11%	3.80%	89.92%
2	0.10%	7.40%	1.27%	4.91%	94.83%
1	0.77%	7.79%	2.29%	5.17%	100.00%

Table 15: Implementing the weighted historical simulation.

Another alternative is the stationary block bootstrap proposed in Politis and Romano [23], that randomly extracts blocks of returns having a variable length. This allows to mimic the stationary properties of the original time series and, at the same time, avoid the strong assumption of independence between successive returns. Another possibility is to combine a parametric model with bootstrap simulation to get the so called Filtered boostrap. This procedure is summarized as follows

1. Assume a particular parametric model for the return series, eventually allowing for time varying volatility, e.g. GARCH effects.

2. Resample from the standardized returns using the estimated conditional volatility.

3. If the volatility is correctly specified, these standardized returns are asymptotically independent and, resorting to the law of large numbers, the bootstrap estimate converge to the true value of the quantity to be estimated.

The estimation of the expected shortfall can be summarized as follows.

Fact 23 (Non parametric estimation of ES) *1. Estimate VaR using empirical quantile. Let $p = (1 - \alpha)$.*

2. Find the number of observations lower than or equal to $-VaR_\alpha$

$$\sum_{i=1}^{T} \mathbf{1}_{\left(r_i \leq -\widehat{VaR}_\alpha\right)}.$$

3. Sum those observations

$$\sum_{i=1}^{T} r_i \mathbf{1}_{\left(r_i \leq -\widehat{VaR}_\alpha\right)}.$$

4. The empirical ES estimate is obtained by taking the ratio of the quantities in steps (2) and (3), ie

$$\widehat{ES}_\alpha (t, t + \Delta) = -\frac{\sum_{i=1}^{T} r_i \mathbf{1}_{\left(r_i \leq -\widehat{VaR}_\alpha\right)}}{\sum_{i=1}^{T} \mathbf{1}_{\left(r_i \leq -\widehat{VaR}_\alpha\right)}}.$$

Example 24 (Expected Shortfall via Historical Simulation) *Let us consider the sorted daily returns in Table 10. The 80% VaR is the order statistics of position 4, ie 4.670%. The 80% Expected Shortfall is therefore*

$$\frac{5.245\% + 5.237\% + 4.927\% + 4.670\%}{4} = 5.02\%,$$

ie the average return loss greater than 4.67% is 5.02%.

4   Assessing a VaR Model

This section introduces the problem of backtesting a VaR model. To do this we first define the concept of VaR violation. Then, using the observed number of violations over a given time horizon, we can statistically evaluate if the adopted VaR model is adequate or not.

Let us suppose that on day t we have estimated a 90% 1-day VaR equal to 4%. If the log return between day t and day $t + 1$, is equal to -4.3%, we have a VaR violation, ie the realized loss (4.3%) is larger than the forecasted one (4%) at the assigned confidence level.

Specifically, the VaR violation $I_{t+\Delta}$ is defined comparing the loss on a portfolio and the reported VaR, $VAR_\alpha(t, t + \Delta)$ as follows

$$I_{t+\Delta}(\alpha) = \begin{cases} 1 & if \quad r(t, t + \Delta) < -VaR_\alpha(t, t + \Delta) \\ 0 & if \quad r(t, t + \Delta) > -VaR_\alpha(t, t + \Delta) \end{cases}$$

If we consider the series I_t for different dates t, the hit sequence, e.g., (0, 0, 1, 0, 0,..., 1), tells us the history of whether or not a loss in excess of the reported VAR has been realized.

Example 25 *Table 16 reports a series of realized daily returns and the VaR forecast of the previous day. Comparing the realized return with the VaR forecast we can detect that there are three VaR violations (coloured in cyan) over 15 days.*

4.1   VaR backtesting

The problem of determining the accuracy of a VAR model can be reduced to the problem of determining whether the hit sequence, $I_t(\alpha)$, $t = 1, ..., n\Delta$, satisfies two key properties.

Day	Daily return	-VAR(90%)	Violation
1	-0.0237	-0.0241	0
2	-0.0129	-0.0241	0
3	-0.0081	-0.0241	0
4	-0.0049	-0.0241	0
5	0.0022	-0.0241	0
6	0.0017	-0.0241	0
7	0.0283	-0.0241	0
8	-0.0267	-0.0241	1
9	-0.0052	-0.0241	0
10	0.0084	-0.0241	0
11	-0.0331	-0.0241	1
12	0.0214	-0.0241	0
13	-0.0271	-0.0241	1
14	0.0215	-0.0241	0
15	-0.0107	-0.0241	0

Table 16: Detecting violations of the $VaR_{90\%}$ comparing realized returns with the VaR forecast.

- Unconditional coverage property. The probability of realizing a loss in excess of the reported VaR must be precisely $(1 - \alpha) \times 100\%$, or, in terms of the previous notation,

$$\Pr_t(I_{t+\Delta}(\alpha) = 1) = 1 - \alpha.$$

If the number of violations occurs more frequently than $(1 - \alpha) \times 100\%$ of the time, this would suggest that the reported VAR measure systematically understates the portfolio's risk. Also a small number of VAR violations is a signal of an overly conservative VAR measure. Both cases suggest that the VaR model is either unable to predict losses or too conservative.

- Independence property. This means that any two elements of the hit sequence, $(I_{t+j\Delta}(\alpha), I_{t+k\Delta}(\alpha))$, $k \neq j$, must be independent random variables. In other words, the past observed VAR violations do not convey any information about whether or not an additional VAR violation will occur in the future.

The above are natural requirements which an accurate VaR system must satisfy. A statistical test for unconditional coverage is the Kupiec

Test, which looks only at the number of violations without examining if they cluster in time, ie if the independence assumption is satisfied. The Kupiec test is also called probability of failure (POF) test.

Let us define $p(j)$ the probability of having j hits given the sample of size n, ie

$$p(j) = \Pr\left(\sum_{i=1}^{n} I_{i\Delta}(\alpha) = j\right).$$

Given that $1 - \alpha$ is the **theoretical** probability of having a violation in each trial, and that $\sum_{i=1}^{n} I_{i\Delta}(\alpha)$ is a binomial (sum of iid Bernoulli's) random variable, it follows that the theoretical probability of having j violations out of n trials is

$$L(j, n, \alpha) = \binom{n}{j} \times (1 - \alpha)^j \times \alpha^{n-j}.$$

Let $1 - \widehat{\alpha}$ be the **observed** violation frequency, i.e.

$$1 - \widehat{\alpha} = \sum_{i=1}^{n} I_{i\Delta}(\alpha) / n.$$

It follows that the **estimated** probability of having j violations out of n trials is

$$L(j, n, \widehat{\alpha}) = \binom{n}{j} \times (1 - \widehat{\alpha})^j \times \widehat{\alpha}^{n-j}.$$

Let us now construct the **log-likelihood ratio**

$$LR_{uc} = -2\ln \frac{L(j, n, \alpha)}{L(j, n, \widehat{\alpha})} = -2\left(j\ln\left(\frac{1-\alpha}{1-\widehat{\alpha}}\right) + (n-j)\ln\left(\frac{\alpha}{\widehat{\alpha}}\right)\right).$$

The above expression is valid if $0 < j < n$. If $j = 0$, then we set $LR_{uc} = -2n\ln(\alpha)$. If instead $j = n$ then $LR_{uc} = -2n\ln(1-\alpha)$. Under the null hypothesis that the VaR model is good, $\widehat{\alpha}$ should not be too different from α, or equivalently the LR_{uc} should have values near 0. For n large and under the null hypothesis H_0: $\widehat{\alpha} = \alpha$, it can be found that the asymptotic distribution of LR_{uc} is chi-square with 1 degree of freedom. Therefore the model is well calibrated if LR_{uc} is less than the critical value of the chi-square distribution with 1 degree of freedom. The critical level is assigned on the basis of the significance level of the test. Alternatively, we can calculate the P-value defined as the probability of getting a sample with a higher LR than the one observed

$$P_{value} = 1 - F_{\chi_1^2}(LR_{uc}),$$

48

where $F_{\chi_1^2}$ is the CDF of a χ^2 random variable with 1 degree of freedom. If the P-value is less than the significance level, we reject, otherwise we accept.

Example 26 *Let us suppose that the sample size used to detect VaR violations is $n = 255$, the VaR confidence level is $\alpha = 0.99$, the probability level of the coverage test is $p = 0.95$, (i.e. the significance level of the test is 0.05). Therefore, the quantile at the 95% level of the chi-square distribution with 1 degree of fredom is equal to 3.84. We cannot reject the model if the number of violations $j = \sum_{i=1}^{n} I_{i\Delta}(\alpha)$ satisfies*

$$-2 \left(j \ln \left(\frac{0.01}{\frac{j}{255}} \right) + (255 - j) \ln \left(\frac{0.99}{1 - \frac{j}{255}} \right) \right) < 3.84,$$

or, with some algebra, if

$$0 < j < 7.$$

Otherwise we have to reject the VaR model at the 5% level. For example, with 10 violations we reject the fact that the model is accurate, indeed

$$LR_{uc} = 12.65 > 3.84,$$

and the P-value of the test is

$$1 - F_{\chi_1^2}(12.65) = 3.8 \times 10^{-4} < 0.05;$$

On the other hands, if we observe 3 violations, the model is deemed to be accurate, indeed

$$LR_{uc} = 0.07591 < 3.84,$$

and the P-value is

$$P_{value} = 1 - F_{\chi_1^2}(0.07591) = 0.78290 > 0.05.$$

In general, the Kupiec's test requires a sample consisting of a large number of data (approximately 10 years of daily data) in order to be able to generate truly reliable results. In addition, the Kupiec test is affected by a low power problem, i.e. it has a relatively high probability of missclassifying inaccurate VaR estimates as *acceptably inaccurate*. The lack of power is related to the difficulty of estimating correctly quantiles at high confidence levels, as we already saw in reference to the estimation risk of the VaR. A high confidence level generates too few exceptions for conducting a reliable test. Indeed this problem is considerably reduced if we lower

the VaR confidence level. A possibility to increase the power of the test is to increase the sample size.

Another limitation of the Kupiec's test is that it focuses only on the number of exceptions generated by a VaR model, without considering the time distribution of these exceptions. In this sense, the test is unconditional: the quality of a model is evaluated independently of its ability to promptly respond to new market conditions. To cope with this problem, an independence test has been proposed by Christoffersen [13], who extended the Kupiec statistic to test that exceptions are serially independent.

All of the mentioned backtests procedures focus on the adequacy of a VaR model at a given confidence level. But, in principle this is not necessary. The unconditional coverage and independence property should hold for any level of confidence. Crnkovic and Drachman [14], Berkowitz [5], Diebold *et al* [16] have all suggested backtests based on multiple VaR levels. For a review of these, see Campbell [10].

5 Introduction to Portfolio Modelling

Aim of this second part is to introduce the main tools for assessing the relevant risk measures, as VaR and ES, at portfolio level. The main issue is how to specify the joint distribution of the log-returns of the portfolio components. In addition, the inclusion of non-linear derivative positions in the portfolio makes it difficult to obtain the portfolio distribution. In this case, Monte Carlo simulation is of great help. Concrete examples from energy markets are also considered.

As we move from exposure at a single asset level to exposure at portfolio level, non trivial issues arise because we need to be able to capture the dependence structure among the portfolio components. A possible solution is the so called top-down approach, i.e. the porfolio P&L distribution is assigned without reference to the portfolio constituents. Then the computation of the portfolio VaR can be done by referring to the approaches previously presented, treating the portfolio return series as a univariate series. The limit of this approach is that it does not allow us to identify the assets contribution to the portfolio exposure: a large loss can occurr at portfolio level, but it is not possible to identify the asset that most contributed to it. In the top down approach the dependence structure among constituents is implicit in the specification of the portfolio distribution. Alternatively, we can adopt a bottom-up approach in which we model the complex interactions at the level of individual instruments and the porfolio exposure is an aggregate of the constituent. Diversification effects are

captured by modelling directly the dependencies among the different assets. Therefore, if the number of commodities is large and the sample size is limited, the model estimation in the bottom-up approach can be quite problematic. In general, the bottom-up approach is feasible in the context of the parametric approach, specifically if we assume that commodity returns are jointly Gaussian, because in this case the dependence structure is entirely captured by the covariance matrix. In this section, we illustrate the bottom-up approach in the non parametric and in the parametric case.

We recall that if we have N commodity positions in the portfolio with weigth w_i and each commodity has a log-return $r_i\,(t, t + \Delta)$, the portfolio log-return $r_p\,(t, t + \Delta)$ is given by

$$r_p\,(t, t + \Delta) = \ln\left(\sum_{i=1}^{N} w_i e^{r_i(t,t+\Delta)}\right) = \ln\left(\mathbf{w}' e^{\mathbf{r}(t,t+\Delta)}\right), \qquad (15)$$

where $\mathbf{r}$ is the $Nx1$ random vector having as components the random log-return of the different commodities and $\mathbf{w}$ is the $N \times 1$ vector containing the weights w_i.

In general the distribution of r_p is not known in closed form, even if the log-returns are jointly Gaussian.

5.1 Historical Simulation

The application of Historical Simulation to a portfolio context is straigth-forward.

1. For each period in the sample, we compute the log-return of the different assets, $r_i\,(t, t + \Delta)$ for all $i = 1, \cdots, N$.

2. Then for each period we compute the portfolio return via (15).

3. We resample from the past the portfolio returns and we build M samples each having the same size T as the original one.

4. Out of each simulated sample we compute the portfolio VaR and then we average the M values to obtain the bootstrap estimate of the VaR.

5. If needed, we can use the simulated VaR values to compute a confidence interval around the bootstrap VaR estimate.

Example 27 (Porfolio P&L via Historical Simulation) *Let us consider three energy products: Europe Brent Spot Price FOB (Dollars per Barrel), New York Harbor Conventional Gasoline Regular Spot Price FOB (Dollars per Gallon) and New York Harbor No. 2 Heating Oil Spot Price FOB (Dollars per Gallon). Their daily returns in August 2015 are given in columns 2-4 of Table 17. In the fifth column, for each day of the sample, we compute the portfolio return according to formula (15), assuming that the weights are equal to 1/3. Then we run the bootstrap simulation randomly sampling over the 20 historical portfolio returns. This is done in the remaining columns. In column 6, we randomly extract integer numbers in the range [1,20] and in the adjacent column we resample the corresponding portfolio returns. This is also done in the last columns of the same Table. In the bottom line, for each simulated sample we compute the portfolio 90% VaR. If we repeat the simulation $M = 500$ times, we obtain M simulated VaR estimates that convey the same information as the original sample, if the i.i.d. assumption is satisfied. By averaging these estimates we obtain the bootstrapped estimate of the VaR. A confidence interval can also be constructed. The results are shown in Table 18. A similar procedure can be followed in order to obtain the boostrap estimate of the ES.*

5.2 The Gaussian covariance approach

In this approach, we assume that the log-returns of the different commodities have a joint normal distribution. We express this using the notation

$$\mathbf{r}(t, t+\Delta) \sim \mathcal{N}(\boldsymbol{\mu}_\Delta, \boldsymbol{\Sigma}_\Delta), \tag{16}$$

where $\boldsymbol{\mu}_\Delta$ is the mean vector of the expected return of the different assets over the period of length Δ,

$$\boldsymbol{\mu}_\Delta = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_i \\ \vdots \\ \mu_N \end{bmatrix} = \begin{bmatrix} \mathbb{E}(r_1(t, t+\Delta)) \\ \vdots \\ \mathbb{E}(r_i(t, t+\Delta)) \\ \vdots \\ \mathbb{E}(r_N(t, t+\Delta)) \end{bmatrix}.$$

$\boldsymbol{\Sigma}_\Delta$ is the $N \times N$ covariance matrix. It has to be symmetric and positive definite[10]. It has on the main diagonal the N variances of the different assets and out of the main diagonal the $N(N-1)/2$ cross-covariances (denoted

---

[10]A matrix $\boldsymbol{\Sigma}$ is positive definite if $\mathbf{x}'\boldsymbol{\Sigma}\mathbf{x} > 0$ for all $\mathbf{x} \in R^N, \mathbf{x} \neq \mathbf{0}$.

t	Brent	Gasoline	Heating Oil	Port. Ret.	Sim. nr. 1			Sim. nr. 500	
					$\mathcal{U}$	$r_{p,\mathcal{U}}$		$\mathcal{U}$	$r_{p,\mathcal{U}}$
1	-0.0083	0.0074	0.0149	0.0047	8	-0.0141	$\cdots$	19	0.0587
2	-0.0008	-0.0524	-0.0099	-0.0210	12	-0.0346	$\cdots$	12	-0.0346
3	-0.0256	-0.0114	0.0078	-0.0097	4	-0.0105	$\cdots$	11	-0.0114
4	-0.0055	-0.0146	-0.0114	-0.0105	16	-0.0092	$\cdots$	1	0.0047
5	0.0159	0.0444	0.0275	0.0293	20	0.0442	$\cdots$	20	0.0442
6	-0.0203	0.0047	-0.0126	-0.0094	3	-0.0097	$\cdots$	2	-0.0210
7	0.0201	0.0339	0.0161	0.0233	5	0.0293	$\cdots$	8	-0.0141
8	-0.0058	-0.0246	-0.0119	-0.0141	4	-0.0105	$\cdots$	1	0.0047
9	-0.0046	-0.0175	-0.0134	-0.0118	8	-0.0141	$\cdots$	7	0.0233
10	-0.0004	-0.0178	0.0078	-0.0035	3	-0.0097	$\cdots$	14	-0.0157
11	-0.0163	-0.0151	-0.0028	-0.0114	2	-0.0210	$\cdots$	19	0.0587
12	-0.0270	-0.0524	-0.0244	-0.0346	19	0.0587	$\cdots$	7	0.0233
13	-0.0026	-0.0149	-0.0168	-0.0114	13	-0.0114	$\cdots$	15	-0.0493
14	-0.0400	0.0123	-0.0194	-0.0157	17	-0.0162	$\cdots$	19	0.0587
15	-0.0527	-0.0467	-0.0486	-0.0493	18	0.0674	$\cdots$	11	-0.0114
16	0.0065	-0.0197	-0.0143	-0.0092	20	0.0442	$\cdots$	16	-0.0092
17	-0.0024	-0.0493	0.0032	-0.0162	2	-0.0210	$\cdots$	11	-0.0114
18	0.0627	0.0548	0.0849	0.0674	11	-0.0114	$\cdots$	8	-0.0141
19	0.0760	0.0402	0.0598	0.0587	14	-0.0157	$\cdots$	16	-0.0092
20	0.0000	0.0800	0.0525	0.0442	10	-0.0035	$\cdots$	11	-0.0114
				90% VaR		**0.0210**	$\cdots$		**0.0346**

Table 17: Historical Simulation of the portfolio distribution. Columns labelled with $\mathcal{U}$ refer to simulated integer uniform numbers in the interval 1-20. Adjacent columns refer to resampled portfolio returns from column 5.

	VaR	VaR 95% Confidence Interval	
Estimate	2.82%	1.38%	4.93%

Table 18: Bootstrap VaR estimate using 500 simulations. A 95% confidence interval aroung the point estimate is also given.

by σ_{ij} with $\sigma_{ij} = \sigma_{ji}$).

$$\boldsymbol{\Sigma}_\Delta = \begin{bmatrix} \sigma_1^2 & & \cdots & & \sigma_{1,N} \\ & \sigma_2^2 & & & \\ \vdots & & \ddots & & \vdots \\ & & & \sigma_{N-1}^2 & \\ \sigma_{N,1} & & \cdots & & \sigma_N^2 \end{bmatrix}$$

where

$$\sigma_i^2 = \mathbb{V}\left(r_i(t, t+\Delta)\right); \sigma_{ij} = \mathbb{C}ov\left(r_i(t, t+\Delta), r_j(t, t+\Delta)\right).$$

The portfolio return is a non-linear function of the component return, see formula (15), so it does not have a Gaussian distribution, even if the components do. Therefore, it is convenient, from the analytical point of view, to replace (15) with the following linear approximation

$$r_p(t, t+\Delta) \simeq \sum_{i=1}^{N} w_i r_i(t) = \mathbf{w}'\mathbf{r}(t). \tag{17}$$

This approximation is in general invalid over long time horizon (e.g. longer than 1 year), or in the case in which volatilities are large or short positions are allowed. Using the approximation (17) and exploiting the properties of the expectation and of the variance, it follows that the portfolio expected return and variance are

$$\mathbb{E}\left(r_p(t, t+\Delta)\right) = \mathbb{E}\left(\mathbf{w}'\mathbf{r}(t, t+\Delta)\right) = \mathbf{w}'\mathbb{E}\left(\mathbf{r}(t, t+\Delta)\right) = \mathbf{w}'\boldsymbol{\mu}_\Delta,$$

and

$$\mathbb{V}\left(r_p(t, t+\Delta)\right) = \mathbb{V}\left(\mathbf{w}'\mathbf{r}(t, t+\Delta)\right) = \mathbf{w}'\mathbb{V}\left(\mathbf{r}(t, t+\Delta)\right)\mathbf{w} = \mathbf{w}'\boldsymbol{\Sigma}_\Delta\mathbf{w}.$$

In addition, we have the following result on the portfolio distribution.

Fact 28 (Portfolio log-return and its distribution) *Under the modelling assumption (16), and using (17), it follows that*

$$r_p(t, t+\Delta) = \mathbf{w}'\mathbf{r}(t, t+\Delta) \sim \mathcal{N}\left(\mathbf{w}'\boldsymbol{\mu}_\Delta, \mathbf{w}'\boldsymbol{\Sigma}_\Delta\mathbf{w}\right),$$

i.e. the portfolio log-return is Gaussian. In addition, for a generic time horizon composed of n periods, and assuming zero serial-correlation, we have

$$r_p(t, t+n\Delta) = \mathbf{w}'\mathbf{r}(t, t+n\Delta) \sim \mathcal{N}\left(\mathbf{w}'\boldsymbol{\mu}_\Delta n, \mathbf{w}'\boldsymbol{\Sigma}_\Delta\mathbf{w}n\right).$$

We can now use the VaR and ES formulae we gave in the univariate case. The portfolio VaR in return terms is

$$VaR_\alpha^r\left(t, t+n\Delta\right) = -\left(\mathbf{w}'\boldsymbol{\mu}_\Delta n + z_{1-\alpha}\sqrt{\mathbf{w}'\boldsymbol{\Sigma}_\Delta \mathbf{w} \times n}\right). \tag{18}$$

The portfolio VaR in dollar terms is

$$VaR_\alpha^{P\&L}\left(t, t+n\Delta\right) = P\left(t\right)\left(1 - e^{-VaR_\alpha^r(t,t+n\Delta)}\right),$$

where $P\left(t\right)$ is the current portfolio value. The portfolio Expected Shortfall is

$$ES_\alpha^r\left(t, t+n\Delta\right) = -\mathbf{w}'\boldsymbol{\mu}_\Delta n + \frac{\phi\left(z_{1-\alpha}\right)}{1-\alpha}\sqrt{\mathbf{w}'\boldsymbol{\Sigma}_\Delta \mathbf{w}} \times \sqrt{n}, \tag{19}$$

$$ES_\alpha^{P\&L}\left(t, t+n\Delta\right) \simeq P\left(t\right) ES_\alpha^r\left(t, t+n\Delta\right), \tag{20}$$

where $\phi\left(z_{1-\alpha}\right)$ is the density of the univariate standard Gaussian random variable.

Example 29 (Computing portfolio VaR) *Let us consider the following inputs*

$$\mathbf{w}' = \left[\begin{array}{ccc} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array}\right],$$
$$\boldsymbol{\mu}_\Delta' = \left[\begin{array}{ccc} 0.01 & 0.02 & 0.04 \end{array}\right],$$
$$\boldsymbol{\Sigma}_\Delta = \left[\begin{array}{ccc} 0.1230 & 0.1290 & 0.1420 \\ 0.1290 & 0.1940 & 0.1670 \\ 0.1420 & 0.1670 & 0.1840 \end{array}\right].$$

We can compute

1. The portfolio expected return

$$\mathbf{w}'\boldsymbol{\mu}_\Delta = \left[\begin{array}{ccc} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array}\right]\left[\begin{array}{c} 0.01 \\ 0.02 \\ 0.04 \end{array}\right] = \frac{0.07}{3}.$$

2. The portfolio variance

$$\mathbf{w}'\boldsymbol{\Sigma}_\Delta \mathbf{w} = \left[\begin{array}{ccc} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array}\right]\left[\begin{array}{ccc} 0.1230 & 0.1290 & 0.1420 \\ 0.1290 & 0.1940 & 0.1670 \\ 0.1420 & 0.1670 & 0.1840 \end{array}\right]\left[\begin{array}{c} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{array}\right] = \frac{1.377}{3^2}.$$

3. *The 10 days portfolio VaR (return terms)*

$$VaR_\alpha^r(t, t + n\Delta) = -\left(0.07 \times 10 + z_{1-\alpha}\sqrt{1.377 \times 10}\right) \times \frac{1}{3}.$$

4. *The 10 days portfolio Expected Shortfall (return terms)*

$$ES_\alpha^r(t, t + n\Delta) = -\left(-0.07 \times n + \frac{e^{-\frac{1}{2}\left(z_{1-\alpha}^2\right)}}{\sqrt{2\pi}(1 - \alpha)}\sqrt{1.377} \times \sqrt{n}\right)\frac{1}{3}.$$

The implementation of the VaR formula requires the estimation of the mean vector and the covariance matrix. As in the univariate case, the simplest approach is to use unconditional estimators, i.e. the sample mean vector and the sample covariance matrix. This is presented in the next example.

Example 30 (Portfolio VaR via the Gaussian Covariance Method) *Let us consider an equally weighted portfolio made of three commodities (Brent, Gasoline and Heating Oil) whose log-returns are given in Table 17. We can estimate the expected returns*

- *for Brent*

$$\widehat{\mu}_{brent} = \frac{1}{20}\sum_{i=1}^{20} r_{brent}(i\Delta) = -0.0016;$$

- *for Gasoline*

$$\widehat{\mu}_{gas} = \frac{1}{20}\sum_{i=1}^{20} r_{gas}(i\Delta) = -0.0029;$$

- *for Heating Oil*

$$\widehat{\mu}_{ho} = \frac{1}{20}\sum_{i=1}^{20} r_{ho}(i\Delta) = 0.0045.$$

In addition, we can estimate the second moments using the corresponding sample estimates

$$\widehat{\mu}_{brent,brent} = \frac{1}{21}\sum_{t=1}^{21} r_{brent}(i\Delta)r_{brent}(i\Delta) = 0.0008,$$

$$\widehat{\mu}_{brent,gas} = \frac{1}{21}\sum_{t=1}^{21} r_{brent}(i\Delta)r_{gas}(i\Delta) = 0.0006,$$

$$\widehat{\mu}_{brent,ho} = \frac{1}{21} \sum_{t=1}^{21} r_{brent}(i\Delta)r_{ho}(i\Delta) = 0.0007,$$

and

$$\widehat{\mu}_{gas,gas} = \frac{1}{21} \sum_{t=1}^{21} r_{gas}(i\Delta)r_{gas}(i\Delta) = 0.0013,$$

$$\widehat{\mu}_{gas,ho} = \frac{1}{20} \sum_{i=1}^{20} r_{gas}(i\Delta)r_{ho}(i\Delta) = 0.0009,$$

$$\widehat{\mu}_{ho,ho} = \frac{1}{21} \sum_{t=i}^{20} r_{ho}(i\Delta)r_{ho}(i\Delta) = 0.0010.$$

The sample variances and covariances are:

$$\widehat{\sigma}^2_{brent} = \widehat{\mu}_{brent,brent} - (\widehat{\mu}_{brent})^2 = 0.0008 - (-0.0016)^2 = 0.0008,$$

$$\widehat{\sigma}^2_{gas} = \widehat{\mu}_{gas,gas} - (\widehat{\mu}_{gas})^2 = 0.0013 - (-0.0029)^2 = 0.0013,$$

$$\widehat{\sigma}^2_{ho} = \widehat{\mu}_{ho,ho} - (\widehat{\mu}_{ho})^2 = 0.0010 - (0.0045)^2 = 0.0010,$$

$$\widehat{\sigma}^2_{brent,gas} = \widehat{\mu}_{brent,gas} - \widehat{\mu}_{brent}\widehat{\mu}_{gas} = 0.0006 - (-0.0016)(-0.0029) = 0.0006$$

$$\widehat{\sigma}^2_{brent,ho} = \widehat{\mu}_{brent,ho} - \widehat{\mu}_{brent}\widehat{\mu}_{ho} = 0.0007 - (-0.0016)(0.0045) = 0.0007,$$

$$\widehat{\sigma}^2_{gas,ho} = \widehat{\mu}_{gas,ho} - \widehat{\mu}_{gas}\widehat{\mu}_{ho} = 0.0009 - (-0.0029)(0.0045) = 0.0009.$$

Notice that in practice, using the means in the above formula does not affect the estimates of the covariances. For this reason, it is common practice to set at zero the estimated daily means.

By collecting sample means, sample variances and covariances we have the estimated sample mean vector and the sample covariance matrix.

Asset	Mean Vector	Covariance Matrix		
Brent	-0.0016	0.0008	0.0006	0.0007
Gasoline	-0.0029	0.0006	0.0013	0.0009
Heat. Oil	0.0045	0.0007	0.0009	0.0010

Table 19: Sample mean vector and sample covariance matrix.

We can now compute

1. *The portfolio expected return*

$$\mathbb{E}(r_p) = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} -0.0016 \\ -0.0029 \\ 0.0045 \end{bmatrix} = 0.$$

2. *The portfolio variance*

$$\mathbb{V}(r_p) = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 0.0008 & 0.0006 & 0.0007 \\ 0.0006 & 0.0013 & 0.0009 \\ 0.0007 & 0.0009 & 0.0010 \end{bmatrix} \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}$$

   so that $\mathbb{V}(r_p) = 0.0008.$

3. *Finally the estimated portfolio VaR at the 95% confidence level is*

$$\widehat{VaR}_{0.95} = -(0 \times 10 - 1.64485 \times \sqrt{0.0008 \times 10}) = 0.1515.$$

   The (estimated) maximum loss on the portfolio over the next 10 days horizon at 95% confidence level is 15%.

5.3   Issues in estimating the covariance matrix

The remarks about the limits of the sample variance observed in the univariate case also apply to the estimation of the covariance matrix using the sample estimator. This is an unconditional estimator.

Unfortunately, it is difficult to find liquid market products that can be of some help in extracting the implied covariances. An example are Crack Spread Options traded on CME since 1992; the underlying asset of these contracts is the spread of futures on the crack spread, which represents the theoretical refining margin between the refined products (gasoline or diesel) and crude oil. They are American-style options, i.e. they allow for early exercise before maturity. Given a pricing model, these options can provide information for the implied correlation between the two underlying, see Caldana and Fusai [8] and Caldana et al. [9] for recent advances in pricing spread and basket options. However, this can be of limited help in a large portfolio including very different energy products.

Another possibility is to extend the EWMA recursion, and its GARCH variations, to the multivariate case. This allows to capture a time-varying effect in the conditional volatility and in the conditional covariances. The extension is relatively simple for the EWMA model, more cumbersome for

GARCH models. We assume that returns are conditionally Gaussian with time-varying covariance matrix

$$\mathbf{r}\left(t, t+\Delta\right)|\mathcal{F}_t \sim \mathcal{N}\left(\mathbf{0}, \mathbf{\Sigma}_\Delta\left(t\right)\right),$$

where the conditional covariance matrix follows the recursion

$$\mathbf{\Sigma}_\Delta\left(t\right) = \lambda\mathbf{\Sigma}_\Delta\left(t-\Delta\right) + \left(1-\lambda\right)\mathbf{r}\left(t-\Delta, t\right)\mathbf{r}'\left(t-\Delta, t\right). \quad (21)$$

Notice that here we have a unique value of λ that applies to all entries of the covariance matrix. This guarantees that the covariance matrix to be positive definite. λ can be estimated again via ML.

Example 31 (The Multivariate EWMA model) *Starting from an initial co-variance matrix*

$$\Sigma_\Delta(0) = \begin{bmatrix} 9 & 8 \\ 8 & 16 \end{bmatrix}$$

and given the sequence of returns

Day	r'	
1	-3	0
2	0	-3
3	1	-2
4	-3	-6

we can perform the following calculations.

- **Step 1**

$$\mathbf{\Sigma}_\Delta(\Delta) = 0.9 \times \begin{bmatrix} 9 & 8 \\ 8 & 16 \end{bmatrix} + 0.1 \times \begin{bmatrix} -3 \\ 0 \end{bmatrix} \begin{bmatrix} -3 & 0 \end{bmatrix} = \begin{bmatrix} 9 & 7.2 \\ 7.2 & 14.4 \end{bmatrix}.$$

- **Step 2**

$$\begin{aligned}
\mathbf{\Sigma}_\Delta(2\Delta) &= 0.9 \times \begin{bmatrix} 9 & 7.2 \\ 7.2 & 14.4 \end{bmatrix} + 0.1 \times \begin{bmatrix} 0 \\ -3 \end{bmatrix} \begin{bmatrix} 0 & -3 \end{bmatrix} \\
&= \begin{bmatrix} 8.1 & 6.48 \\ 6.48 & 13.86 \end{bmatrix}.
\end{aligned}$$

- **Step 3**

$$\begin{aligned}
\mathbf{\Sigma}_\Delta(3\Delta) &= 0.9 \times \begin{bmatrix} 8.1 & 6.48 \\ 6.48 & 13.86 \end{bmatrix} + 0.1 \times \begin{bmatrix} 1 \\ -2 \end{bmatrix} \begin{bmatrix} 1 & -2 \end{bmatrix} \\
&= \begin{bmatrix} 7.39 & 5.632 \\ 5.632 & 12.874 \end{bmatrix}.
\end{aligned}$$

- **Step 4**

$$\boldsymbol{\Sigma}_\Delta(4\Delta) = 0.9 \times \begin{bmatrix} 7.39 & 5.632 \\ 5.632 & 12.874 \end{bmatrix} + 0.1 \times \begin{bmatrix} -3 \\ -6 \end{bmatrix} \begin{bmatrix} -3 & -6 \end{bmatrix}$$

$$= \begin{bmatrix} 7.551 & 6.8688 \\ 6.8688 & 15.1866 \end{bmatrix}.$$

In conclusion, we have

$$\boldsymbol{r}(4\Delta, 5\Delta) \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 7.551 & 6.8688 \\ 6.8688 & 15.1866 \end{bmatrix}\right).$$

So if we are interested in the 1-period VaR, we have to compute

$$\widehat{VaR}_\alpha(4\Delta, 5\Delta) = -z_{1-\alpha}\sqrt{\mathbf{w}'\boldsymbol{\Sigma}_\Delta(4\Delta)\mathbf{w}}.$$

VaR for longer horizons can be estimated using Monte Carlo simulation.

6 Hot spot: Measuring the risk contribution of the portfolio components

Given a portfolio composition $\mathbf{w}$ and the corresponding risk measure $\rho_\alpha(w)$, for example either the VaR or the ES, what is the change in $\rho(\mathbf{w})$ given a change in $\Delta\mathbf{w}$ in the current portfolio allocation? We can define three metrics that help us to identify the true sources of risk in our portfolio: Marginal Risk (MRisk), Incremental Risk (IRisk) and Component Risk (CRisk).

Let $\rho_\alpha(w)$ be our preferred risk measure given a portfolio allocation $\mathbf{w}$. We have the following definitions.

Definition 32 (`Marginal Risk` **)** *MRisk is a N $\times$ 1 vector collecting the change in the portfolio risk from an additional dollar of exposure to a given component. It is defined as the gradient vector, i.e. the vector collecting the partial derivatives of the risk-measure with respect to the component portfolio*

$$\mathbf{MRisk} = \frac{\partial\rho_\alpha(\mathbf{w})}{\partial\mathbf{w}}.$$

Definition 33 (`Incremental Risk` **)** *IRisk is a number that represents the change in the portfolio risk due to an additional change of exposure to a given component. It is defined as*

$$IRisk = \rho_\alpha(\mathbf{w} + \Delta\mathbf{w}) - \rho_\alpha(\mathbf{w});$$

using the first-order Taylor's formula, it can be approximated via the vector multiplication

$$IRISK \sim \mathbf{MRisk'}\Delta\mathbf{w}.$$

Definition 34 (`Component Risk`) *CRisk is a $N \times 1$ vector whose elements indicate the contribution of each portfolio's component to the overall risk. It is defined as*

$$\mathbf{CRisk} = \mathbf{w}. * \mathbf{MRisk}.$$

The sum of the vector components is equal to the portfolio risk[11]

$$\rho = \mathbf{1'CRisk} = \sum_{i=1}^{n} CRisk_i.$$

6.1 Hot Spot in the Gaussian case

If we assume that portfolio components and portfolio returns have a joint Gaussian distribution we have the following result.

Fact 35 (Computing MVaR, IVaR, CVaR in the Gaussian setting) *If we assume that returns are jointly Gaussian, from formula (18) we obtain*

$$\mathbf{MVaR} = \frac{\partial VaR^r_\alpha(t, t + n\Delta)}{\partial \mathbf{w}} = -\boldsymbol{\mu}_\Delta n - z_{1-\alpha} \frac{\boldsymbol{\Sigma}_\Delta \mathbf{w}}{\sqrt{\mathbf{w'}\boldsymbol{\Sigma}_\Delta \mathbf{w}}} \sqrt{n}.$$

Consequently, the IVaR is

$$IVaR_\alpha(\mathbf{w}) = \sum_{i=1}^{N} \underbrace{\Delta w_i \underbrace{MVaR_i}_{\textit{Marginal VaR}}}_{\textit{Incremental VaR due to asset i}}.$$

The CVaR in absolute terms is

$$\mathbf{CVaR} = \begin{bmatrix} -w_1\mu_1 n - w_1 \frac{z_{1-\alpha}\sqrt{n}cov(r_1, r_p)}{\sqrt{\mathbf{w'}\boldsymbol{\Sigma}_\Delta \mathbf{w}}} \\ ... \\ -w_N\mu_N n - w_N \frac{z_{1-\alpha}\sqrt{n}cov(r_N, r_p)}{\sqrt{\mathbf{w'}\boldsymbol{\Sigma}_\Delta \mathbf{w}}} \end{bmatrix},$$

and in percentage terms is

$$CVaR_i\% = \frac{CVaR_i}{VaR_\alpha(\mathbf{w})}.$$

---

[11]This last fact holds if the risk measure is a homogeneous risk function. This property is satisfied by VaR and ES.

It is easy to verify that the sum of Component VaRs returns the portfolio VaR, i.e.

$$VaR_\alpha = \sum_{i=1}^{N} w_i \underbrace{\underbrace{MVaR_i}_{Marginal\ VaR}}_{Component\ VaR} .$$

Indeed

$$VaR_\alpha^r(t, t+\Delta) = \mathbf{1}'\mathbf{CVaR} = \mathbf{w}'\mathbf{MVaR} = -\mathbf{w}'\mu_\Delta n - z_{1-\alpha}\frac{\mathbf{w}'\Sigma_\Delta\mathbf{w}}{\sqrt{\mathbf{w}'\Sigma_\Delta\mathbf{w}}}\sqrt{n}.$$

Notice in the above expressions the appearance of the term $\Sigma_\Delta\mathbf{w}$ which represents the covariance of the portfolio return with the different components[12]. Therefore, in the Gaussian framework we have the simplifying result that the risk contribution of an asset is measured by its covariance with the portfolio. Being the covariance a measure of linear dependence, non linear dependences cannot be captured in the Gaussian framework.

Example 36 (Marginal, Incremental and Component VaR in the Gaussian framework)
Let assume a zero-mean vector and a covariance matrix as estimated in Example 30

$$\Sigma_\Delta = \begin{bmatrix} 0.000847 & 0.000596 & 0.000744 \\ 0.000596 & 0.001335 & 0.000902 \\ 0.000744 & 0.000902 & 0.000953 \end{bmatrix}.$$

Let us consider the case of a trading company which is investing in the following portfolio

$$\mathbf{w}' = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \end{bmatrix}.$$

We have

$$\Sigma_\Delta\mathbf{w} = \begin{bmatrix} Cov(\tilde{r}_p, \tilde{r}_1) \\ Cov(\tilde{r}_p, \tilde{r}_2) \\ Cov(\tilde{r}_p, \tilde{r}_3) \end{bmatrix} = \begin{bmatrix} 0.000746 \\ 0.000893 \\ 0.000831 \end{bmatrix}.$$

The portfolio variance is

$$\sigma_p^2 = \mathbf{w}'\Sigma_\Delta\mathbf{w} = 0.000810.$$

---

[12]Indeed

$$Cov(\tilde{r}_p, \tilde{r}_j) = Cov\left(\sum_{i=1}^{n} w_i\tilde{r}_i, \tilde{r}_j\right) = \sum_{i=1}^{n} w_i Cov(\tilde{r}_i, \tilde{r}_j) = \sum_{i=1}^{n} w_i\sigma_{ij},$$

and in vector form we have $Cov(\tilde{r}_p, \tilde{\mathbf{r}}) = \Sigma_\Delta\mathbf{w}$.

The portfolio VaR is

$$VaR_\alpha^{\mathbf{w}} = -z_{1-\alpha}\sqrt{n}\sqrt{0.000810} = -z_{1-\alpha}\sqrt{n} \times 0.028452.$$

The Marginal VaR is

$$\mathbf{MVaR} = -z_{1-\alpha}\sqrt{n}\begin{bmatrix} 0.026232 \\ 0.031398 \\ 0.029223 \end{bmatrix}.$$

The Component VaR is

$$\mathbf{CVaR} = -z_{1-\alpha}\sqrt{n}\begin{bmatrix} \frac{1}{2} \times 0.026232 \\ \frac{1}{3} \times 0.031398 \\ \frac{1}{6} \times 0.029223 \end{bmatrix} = -z_{1-\alpha}\sqrt{n}\begin{bmatrix} 0.013116 \\ 0.010466 \\ 0.004870 \end{bmatrix}.$$

The sum of the Component VaR is

$$-z_{1-\alpha}\sqrt{n}\left(0.013116 + 0.010466 + 0.004870\right) = -z_{1-\alpha}\sqrt{n} \times 0.02845,$$

i.e. exactly the portfolio VaR. So we can use the CVaR to measure the percentage contribution of each asset to the overall portfolio risk

$$\mathbf{CVaR\%} = \frac{\mathbf{CVaR}}{VaR} = \begin{bmatrix} 46.10\% \\ 36.78\% \\ 17.12\% \end{bmatrix}.$$

    The issue is now how to rebalance the portfolio among the different commodities in order to reduce the overall VaR. To understand this, we have to examine the MVaR vector: it signals that the largest VaR reduction can be achieved by reducing the investment in the second commodity (Gasoline) and increasing the investment in the first one (Brent). Therefore, if we decide to increase by 5% the weigth allocated to Brent to 55%, and to decrease by the same amount the weigth allocated to Gasoline, we can use the IVaR to estimate the effect on the portfolio VaR. We have

$$IVaR = \mathbf{MVaR}'\begin{bmatrix} \frac{5}{100} \\ -\frac{5}{100} \\ 0 \end{bmatrix} = -z_{1-\alpha}\sqrt{n} \times (-0.00026)$$

    This means that the effect of this trade on the VaR of the trading firm's portfolio can be approximately estimated as

$$VaR_\alpha^{\mathbf{w}+\Delta\mathbf{w}} \sim VaR_\alpha^{\mathbf{w}} - z_{1-\alpha}\sqrt{n} \times (-0.00026),$$

i.e. we can reduce the portfolio VaR (recall that $z_{1-\alpha} < 0$). This procedure is much faster to implement because the MVaR vector is a by-product of the initial VaR computation, and does not require a new estimate of the P&L distribution at the given new portfolio composition.

Another consideration arises from the CVaR vector. Brent has the largest CVaR, mainly due to the large weight allocated to it. Therefore, if the firm needs to decide which commodity exposure to hedge, the best strategy is to hedge the Brent exposure. By hedging it, the trading firm can reduce the portfolio VaR by 46.10%.

In conclusion, IVaR can be relevant to understand the effect on the portfolio VaR of a rebalance of the weight allocated to the different assets. CVaR is useful to understand which position to hedge, so that we can achieve the largest VaR reduction. Finally, we notice that these metrics have been computed using only first order approximation.

6.2   Hot Spot in the non Gaussian case

If we need to compute portfolio risk contribution when returns are not jointly Gaussian, or when we have non-linear products in the portfolio, so that the portfolio distribution is not Gaussian, we have to resort to the following results for the VaR and ES.

Fact 37 (Marginal VaR & VaR decomposition) *Let r_p to be the portfolio return. The Marginal VaR for a general distribution is computed according to*

$$MVaR_i = \frac{\partial VaR_\alpha}{\partial w_i} = -\mathbb{E}\left(r_i \,\middle|\, r_p = -VaR_\alpha\right), i = 1, ..., N.$$

The following VaR decomposition

$$VaR_\alpha = \sum_{i=1}^{N} w_i \underbrace{\underbrace{\mathbb{E}\left(r_i \,\middle|\, r_p = -VaR_\alpha\right)}_{\text{Marginal VaR}}}_{\text{Component VaR}}.$$

still holds.

Fact 38 (Marginal Expected Shortfall & ES decomposition) *The marginal contribution to ES of a single asset is given by*

$$\frac{\partial ES_\alpha}{\partial w_i} = -\mathbb{E}\left(r_i \,\middle|\, r_p < -VaR_\alpha\right). \tag{22}$$

We can decompose the ES as

$$ES_\alpha\left(\mathbf{w}\right) = \sum_{i=1}^{N} w_i \underbrace{\underbrace{\mathbb{E}\left(r_i \, \middle| \, r_p < -VaR_\alpha\right)}_{Marginal\ ES}}_{Component\ ES}.$$

In other words, the MVaR for a given asset is the asset conditional expected return given that the portfolio loss is at VaR level. In practice, it can be shown that these formula agree with those given in the previous subsection where asset and portfolio returns are assumed jointly Gaussian. In the non Gaussian case, the computation is performed via Monte Carlo simulation. Given that, mainly for large values of α, as it is very unlikely to find a significant number of instances of the event $r_p = -VaR_\alpha$, the computation of the conditional expectation via Monte Carlo simulation requires a large number of simulations. In practice, the event $r_p = -VaR_\alpha$ is replaced with

$$-VaR_\alpha - \epsilon < r_p < -VaR_\alpha + \epsilon,$$

where ϵ is a small number, i.e. we compute the (conditional) expectation using a small interval around -VaR. Consequently

$$MVaR_i \sim -\mathbb{E}\left(r_i \, \middle| \, -VaR_\alpha - \epsilon < r_p < -VaR_\alpha + \epsilon\right).$$

Example 39 *Let us consider the following portfolio weights*

$$\mathbf{w\prime} = \begin{bmatrix} 20\% & 30\% & 10\% & 5\% & 35\% \end{bmatrix}$$

Let us also suppose that the portfolio 90% VaR and ES are respectively

$$VaR_{0.90} = 1.99\%; ES_{0.90} = 2.81\%$$

In the first column of Table 20 we consider portfolio returns sorted in increasing order and we focus the attention on portfolio returns near to -1.99%, the VaR level. The remaining columns refer to the returns of the different assets. In order to estimate the MVaR, we compute the expected return of each asset conditional on the portfolio VaR falling in the interval $-VaR_\alpha - \epsilon < r_p < -VaR_\alpha + \epsilon$, where ϵ is a small number. In our numerical example, we set $\epsilon = 0.001$, so that we have to compute the expected returns on the different stocks conditional on the portfolio return falling in the interval

$$-2.09\% < r_p < -1.89\%.$$

r_p	r_1	r_2	r_3	r_4	r_5
. . .	. . .	. . .	. . .	. . .	. . .
-2.09%	-1.91%	-2.10%	-1.85%	-2.14%	-2.25%
-2.09%	-2.11%	-2.69%	-1.47%	-1.81%	-1.77%
-2.08%	-2.85%	-1.64%	-1.94%	-1.45%	-2.16%
-2.08%	-1.71%	-2.64%	-1.52%	-2.28%	-1.94%
-2.06%	-1.45%	-3.91%	-0.63%	0.02%	-1.55%
-2.02%	-1.82%	-1.47%	-1.84%	-2.36%	-2.62%
-2.02%	-1.71%	-2.78%	-1.87%	-2.28%	-1.56%
-2.01%	-1.93%	-1.98%	-1.51%	-2.02%	-2.22%
-2.01%	-1.59%	-2.34%	-2.30%	-2.05%	-1.87%
-2.00%	-2.72%	-2.07%	-2.22%	-2.59%	-1.40%
-2.00%	-1.80%	-3.59%	-1.90%	-1.05%	-0.94%
-2.00%	-2.62%	-0.95%	-2.58%	-2.97%	-2.24%
-1.99%	-1.66%	-2.82%	-1.57%	-0.93%	-1.75%
-1.98%	-2.58%	-1.62%	-1.54%	-1.27%	-2.19%
-1.98%	-2.52%	-1.18%	-2.53%	-1.52%	-2.28%
-1.93%	-2.03%	-3.21%	-1.56%	-1.50%	-0.94%
-1.92%	-2.09%	-1.98%	-2.45%	-2.82%	-1.50%
-1.92%	-1.82%	-1.71%	-2.29%	-1.74%	-2.09%
-1.90%	-2.62%	-1.72%	-1.91%	-2.17%	-1.61%
-1.88%	-2.96%	-1.13%	-2.21%	-2.66%	-1.70%
-1.86%	-0.76%	-4.20%	-1.78%	-0.70%	-0.71%
. . .	. . .	. . .	. . .	. . .	. . .

Table 20: In the first column we have the sorted simulated portfolio returns. In the other columns we have the returns of the different assets. Gray cells refer to the simulations for which the portfolio return falls inside the small interval around the VaR level. The corresponding returns of the other assets are averaged in order to compute the Marginal VaRs.

Asset	1	2	3	4	5
MVaR	2.09%	2.21%	1.89%	1.82%	1.82%
w	20%	30%	10%	5%	35%
CVaR	0.42%	0.66%	0.19%	0.09%	0.64%
CVaR %	21%	33%	9%	5%	32%

Table 21: MVaR and CVA for the different assets given simulated returns in Table 20.

For example, for stock 1, we have

$$MVaR_1 = \frac{2.85\% + 1.71\% + 1.45\% + \cdots + 2.09\% + 1.82\% + 2.62\%}{17} = 2.09\%.$$

Similar computations can be done for the remaining assets. Results are given in Table 21. Notice that

$$\sum_{i=1}^{5} CVaR_i = (0.42 + 0.66 + 0.19 + 0.09 + 0.64)\% = 1.997\% \sim VaR = 1.985\%.$$

In addition, we can consider the effect on the portfolio VaR of a change in the portfolio weights. Let us suppose we want to reduce the exposure to assets 1, 2 and 5 and to increase the exposure to stock 3 and 4 (the choice is related to the CVaR% numbers that show that 1, 2 and 5 are the riskiest assets in our portfolio.) Let us consider the following vector of changes to the different components of the portfolio

$$\Delta w = \begin{bmatrix} -5\% & -5\% & +5\% & +10\% & -5\% \end{bmatrix}$$

We can compute the change in the portfolio VaR computing the Incremental VaR. We have

$$
\begin{aligned}
IVaR &= (-5\% \times 2.09 - 5\% \times 2.21 + 5\% \times 1.89 + 10\% \times 1.82 - 5\% \times 1.82)\,\% \\
&= -0.03\%,
\end{aligned}
$$

i.e. we can reduce the portfolio VaR to 1.99%-0.03%=1.96% (larger reductions are possible with larger changes).

7 Estimating VaR for derivative positions

Different approaches are available to measure the exposure of a portfolio including derivative positions. It is often the case that in commodity portfolios we have very complex derivatives, whose pricing can be very time

	Incremental Risk				
Asset	1	2	3	4	5
$\Delta \mathbf{w}$	-5%	-5%	5%	10%	-5%
MVaR	2.09%	2.21%	1.89%	1.82%	1.82%
$\Delta \mathbf{w_i} \times \mathbf{MVaR_i}$	-0.10%	-0.11%	0.09%	0.18%	-0.09%
IVaR	**-0.03%**				

Table 22: Computing Incremental Risk. -0.03% is obtained by computing $\sum_{i=1}^{5} \Delta w_i \times MVaR_i$.

consuming. The main trade off is between computational effort and accuracy. Therefore, the two most common approaches are full revaluation of the derivative positions or delta-gamma approximations. The first consists in repricing the derivative positions in the simulated scenarios. The second approach consists in approximating the derivative pricing function using a quadratic function of the risk drivers. Sometimes, linear approximations are also used but in general they turn out to be very inaccurate.

7.1 Full Revaluation

VaR can be estimated by combining Monte Carlo simulation with full revaluation as follows. Let $C(P, t)$ be the pricing function that relates the risk factor P to the derivative price at time t.

1. We price the derivative position using the current value of the market factors. This requires the computation of $C(P(t), t)$.

2. We simulate log-returns via our preferred model, either parametric Gaussian distribution or Historical Simulation or some other more advanced model. So we obtain M simulated scenarios

$$r^i(t, t + n\Delta)$$

where $i = 1, ..., M$ refers to the simulation number.

3. We obtain the simulated risk-factor price at the VaR horizon $t + n\Delta$

$$P^i(t + n\Delta) = P(t) \times e^{r^i(t, t+n\Delta)}.$$

4. We revaluate the derivative position at the time horizon under each simulated scenario

$$C_{t+n\Delta}^i = C\left(P^i(t + n\Delta), t + n\Delta\right).$$

5. We compute the M simulated P&L on the derivative position

$$P\&L^i = C^i_{t+n\Delta} - C\left(P(t),t\right).$$

6. The VaR is calculated as the appropriate quantile of the M simulated *P&L*.

Notice that the model used for pricing the derivatives can be different from the one used for simulating the risk factor. In particular, the one used for pricing has to be an arbitrage-free model and specified under the so called risk-neutral measure. The model used in scenario generation has to be specified under the so called real world measure. If the VaR horizon is short, such as a few days or weeks, this distinction is negligible for most markets, except those characterized by extremely volatile returns or severe price jumps.

Example 40 (VaR computation via MC simulation and full revaluation)
Let us consider a plain vanilla call option with strike 300. Let us assume that the underlying annualized volatility is 25%, the option time to maturity is 0.33 years (4 months), the interest rate is 8% and the convenience yield is 3%. The current option premium, according to the Black-Scholes formula, is 22.468. We are interested in the 10-days 95% VaR. Let us assume that percentage prices change is distributed according to a Gaussian distribution with zero mean and 25% annualized volatility, so that

$$r^i(t, t + n\Delta) = 0.25 \times \sqrt{n\Delta} \times \epsilon^i(t),$$

where ϵ^i refers to a standard Gaussian random variable.

In Table 23 we report: in the first column the simulation number, in the second column the simulated standard Gaussian random variable, in the third column the simulated cumulative return, in the fourth column the simulated price, in the fifth column the simulated call option price (taking into account that the option has a maturity shorter of 10 days), and in the last column the simulated derivative P&L. For example, with reference to the first simulation.
Simulation step. *The simulated asset return is*

$$r^1(t, T) = 0 \times \frac{10}{250} + 0.25 \times \sqrt{\frac{10}{250}} \times 0.2314 = 0.0116,$$

so that the simulated underlying price is

$$P^{(1)}\left(t + \frac{10}{250}\right) = 305e^{0.0116} = 308.5493$$

Sim. i	ϵ^i	$r^i(t,T)$	$P^i(T)$	$C(P^i(T);T)$	$P\&L^i$
1	0.2314	0.0116	308.5493	23.3629	0.8949
$\cdots$	-0.3198	-0.0160	300.1618	18.2836	-4.1844
$\cdots$	-0.8860	-0.0443	291.7836	13.8803	-8.5878
5000	0.8892	0.0445	318.8661	30.4500	7.9820

Table 23: Monte Carlo and full revaluation.

Revaluation Step. *We reprice the call option using the Black-Scholes formula, given that the underlying price is now at 308.5493 and the option time to maturity is $0.33 - \frac{10}{250}$. The simulated option price in 10 days turns out to be equal to*

$$C\left(P^{(1)}\left(t + \frac{10}{250}\right), t + \frac{10}{250}\right) = 23.3629,$$

so that the option profit and loss is

$$P\&L^{(1)} = 23.3629 - 22.4681 = 0.8949.$$

We repeat the simulation 5000 times. The simulated P&L of the derivative position is given in Figure 19. Using the simulated P&L distribution we can compute the desired VaR measure via the sample quantile of the simulated P&L. The 10 days VaR and ES for different confidence levels are given in Table 24. In this Table we also give the exact VaR of the option position. Indeed, the call price being an increasing function of the underlying price, the exact VaR computation requires only two evaluations: (1.) pricing the option at the current market situations; (2.) repricing the option at the VaR horizon, assuming a loss in the underlying stock equal to the desired VaR level. The option VaR is then computed taking the difference of (1) and (2). We can see that the Monte Carlo estimate agrees quite well with the exact one.

α	90	95	99
MC+ Full Rev. VaR	11.14	13.24	16.34
Exact VaR	11.19	13.25	16.36

Table 24: VaR for a 10 days horizon on a long position on a call option.

The main limit of the above procedure is that the revaluation step can be very costly, especially for complex derivatives in the portfolio. A possible solution is to replace the repricing step by a linear or a quadratic
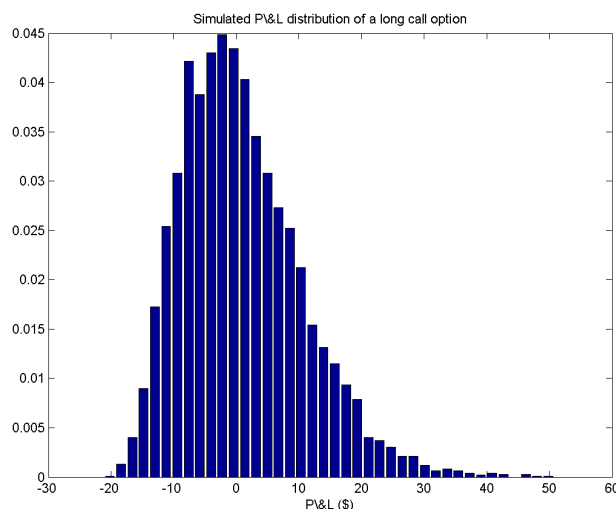
Figure 19: Simulated P&L of the derivative position.

approximation, according to the Taylor's formula. This allows a fast computation of the the derivative price change. This is discussed in the next section.

7.2 VaR and Derivatives: Taylor's Formula & Monte Carlo simulation

The Taylor's formula allows us to approximate the derivative price change according to a polynomial function. In general, a first order (so called delta approximation) or a second order (so called delta-gamma approximation) polynomial are used. Provided we know the distribution of the risk-factors, the Taylor's formula completely defines the distribution of the portfolio P&L and allows us to considerably speed up the repricing step.

If we have several risk factors, it is convenient to denote $C = C(\mathbf{P}, t)$ to be the pricing function of the derivative position with respect to the N risk factors appearing in the vector $\mathbf{P}$, and time. Then we compute

- Theta, i.e. the derivative with respect to time

$$\Theta = \frac{\partial C}{\partial t},$$

- All Delta exposures in the vector called the **dollar gradient vector**

$$\nabla \mathbf{C} = [\Delta_1 P_1, \cdots, \Delta_N P_N]$$

71

where

$$\Delta_i = \frac{\partial C}{\partial P_i}$$

- All Gammas exposures in a matrix called the **dollar-Hessian matrix**

$$\mathbf{H}_C = \begin{bmatrix} \Gamma_{11}P_1^2 & \Gamma_{12}P_1P_2 & \cdots & \Gamma_{1N}P_1P_N \\ \Gamma_{21}P_1P_2 & \Gamma_{22}P_2^2 & \cdots & \Gamma_{2N}P_2P_N \\ \cdots & \cdots & \cdots & \cdots \\ \Gamma_{N1}P_1P_N & \Gamma_{N2}P_NP_2 & \cdots & \Gamma_{NN}P_N^2 \end{bmatrix}$$

where

$$\Gamma_{i,j} = \frac{\partial C}{\partial P_i \partial P_j}, i,j = 1 \cdots, N.$$

- The portfolio P&L is estimated using the second order Taylor formula

$$P\&L = \Theta \Delta t + \nabla \mathbf{C} \frac{d\mathbf{P}}{\mathbf{P}} + \frac{1}{2} \left(\frac{d\mathbf{P}}{\mathbf{P}} \right)' \mathbf{H}_C \frac{d\mathbf{P}}{\mathbf{P}}.$$

This procedure is called Delta-Gamma method. If in the above approximation we consider only first order derivatives, we say we are using the Delta approximation.

Example 41 (Univariate Delta-Gamma Approximation) *Let us consider a call option with time to maturity of 1 year and strike at 300. Further, let us assume the underlying commodity has spot price 305 and volatility 25%. The interest rate is set at 8% and the convenience yield is 3%. If we adopt as pricing model the Black-Scholes model, and assume that the only risk-factor is the percentage change in the underlying price, we can compute the following Greeks*

$$\Theta = -20.2030, \Delta = 0.63287, \Gamma = 0.00470.$$

Consequently, we can approximate the option P&L according to

$$\Delta C(P(t), t)$$

$$= -20.2030 \times dt + 0.63287 \times 305 \times \frac{dP(t)}{P(t)} + \frac{1}{2} \times 0.00470 \times 305^2 \times \left(\frac{dP(t)}{P(t)} \right)^2.$$

Then, we can replace in the above expression the simulated values of $\frac{dP(t)}{P(t)}$ and the simulation of the P&L is immediate. For example, if the simulated percentage price variation is -1%, the simulated P&L is

$$\Delta C(P(t), t) = -20.2030 \times dt + 0.63287 \times 305 \times (-0.01) + \frac{1}{2} \times 0.00470 \times 305^2 \times (-0.01)^2.$$

Example 42 (Multivariate Delta-Gamma Approximation) *Let us assume that the dollar gradient vector is given by:*

$$\nabla \mathbf{C} = [2, 4, -3],$$

the dollar-Hessian matrix is

$$\mathbf{H}_C = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 3 & -2 \end{bmatrix},$$

and the portfolio theta is -0.25. Given a 3 days horizon and a percentage shock to the risk factors equal to

$$\frac{\mathbf{dP}}{\mathbf{P}} = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix},$$

we can estimate the portfolio P&L as

$$P\&L = -0.25 \times \frac{3}{250} + [2, 4, -3] \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix} + \frac{1}{2} [1, -3, 2] \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 3 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}$$

$$= -0.003 - 16 + \frac{1}{2}(-29)$$

$$= -30.503.$$

A Delta approximation is obtained neglecting the second order term, i.e.

$$P\&L = -0.25 \times \frac{3}{250} + [2, 4, -3] \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix} = -0.003 + (2 \times 1 + 4 \times (-3) - 3 \times 2)$$

$$= -16.003.$$

Example 43 (Montecarlo VaR and Delta-Gamma Approximation (univariate case)) *Let us consider the data of Example 40, and the following quantities*

Option Price	*22.4680*
Option Delta	*0.61258*
Option Theta	*-32.46252*
Option Gamma	*0.00857*

In Table 25 we report the results of different Monte Carlo simulations of the derivative P%L.

Sim. i	ϵ^i	$r^i(t,T)$	$\frac{dP^i(t)}{P^i(t)}$	Delta Appr.	Delta-Gamma App.	Full Rev.
1	0.2314	0.0116	0.0116	0.8757	0.9297	0.8949
2	-0.3198	-0.0160	-0.0159	-4.2623	-4.1619	-4.1844
$\cdots$	-0.8860	-0.0443	-0.0433	-9.3946	-8.6459	-8.5878
5000	0.8892	0.0445	0.0455	7.1956	8.0196	7.9820
$VaR_{95\%}$				16.029	13.551	13.242

Table 25: Simulating the derivative P&L using first order (Delta) approximation, second order (Delta-Gamma) approximation and full revaluation. Last row provides the VaR of the derivative position. The exact VaR on the call option is 13.255.

With reference to the first simulation we have that the Delta-Gamma approximation to the option profit and loss is

$$P\&L^{(1)} = -32.4625 \times \frac{10}{250} + 0.61258 \times 305 \times 0.0116 + 0.5 \times 0.00857 \times 305^2 \times 0.0116^2$$
$$= 0.9297.$$

The first-order (Delta) approximation is simply

$$P\&L^{(1)} = -32.4625 \times \frac{10}{250} + 0.61258 \times 305 \times 0.0116 = 0.8757.$$

In conclusion, in order to estimate the portfolio VaR, we proceed as follows

1. Pre-compute Theta, Gradient and Hessian.

2. Simulate M scenarios of the risk factors according to a multivariate risk model
$$\frac{\mathbf{dP^i}}{\mathbf{P}}, i = 1, \cdots, M.$$

3. In each scenario, revaluate the portfolio according to
$$P\&L^i = \Theta\Delta t + \nabla\mathbf{C}\frac{\mathbf{dP^i}}{\mathbf{P}} + \frac{1}{2}\left(\frac{\mathbf{dP^i}}{\mathbf{P}}\right)'\mathbf{H}_C\frac{\mathbf{dP^i}}{\mathbf{P}}.$$

4. Given the M simulated P&L's, $P\&L^1, \cdots, P\&L^M$, we estimate the VaR using the empirical quantile at the desired confidence level.

Example 44 (Multivariate Delta-Gamma Approximation) *Let us consider a portfolio containing options on three commodities (Brent, Gasoline and Heating Oil). The characteristics of these options are given in Table 26. Notice that those are options on futures, so they are priced according to the Black formula, rather than to the Black-Scholes formula. Given the market information, we can compute*

Commodity	Brent	Gasoline	Heating Oil
Trade Date	29/9/15	29/9/15	29/9/15
Type	Call	Put	Call
Option Expiry	Nov, 1st 2016	Nov. 24th 2015	Jan. 26th 2016
Days to Expiry	105	57	120
Time to Maturity (years)	0.2877	0.1562	0.3288
Futures Price	49.41	1.3052	1.5609
Strike	45.5	0.92	1.66
Implied Volatility	41.35%	45%	33.45%
US LIBOR	0.3266%	0.2591%	0.3266%
Disc. Factor	0.99905	0.99959	0.99891
Premium	6.42383	0.00180	0.07966
Delta	0.68466	-0.01990	0.41053
Gamma	0.03237	0.20772	1.29785
Theta	-6.75671	-0.03583	-0.17691
Units	3	4000	150

Table 26: Market Prices and Greeks of options on three commodites. Option premiums and Greeks are computed using the Black formula. Trading date: Sept. 28, 2015.

the portfolio Greeks

- *Theta of the portfolio is*

  $$\Theta = -6.75671 \times 3 - 0.03583 \times 4000 - 0.17691 \times 150 = -190.11357.$$

- *The Dollar Gradient Vector is*

  $$\nabla \mathbf{C} = \begin{bmatrix} 101.48743 & \text{-}103.90143 & 96.11865 \end{bmatrix},$$

  where the first component has been computed as

  $$0.68466 \times 3 \times 49.41 = 101.48743.$$

- *The Dollar Gamma Matrix is*

$$\mathbf{H} = \begin{bmatrix} 237.10046 & 0 & 0 \\ 0 & 1415.44917 & 0 \\ 0 & 0 & 474.31453 \end{bmatrix}$$

where $237.10046 = 0.03237 \times 3 \times (49.41)^2$.

Covariance Matrix			
Brent	0.000847	0.000596	0.000744
Gasoline	0.000596	0.001335	0.000902
Heating Oil	0.000744	0.000902	0.000953

Table 27: Covariance matrix of daily log-returns.

Let us suppose that log-returns of the three commodities are jointly Gaussian with zero mean and covariance matrix as reported in Table 27. In addition, let us suppose that the VaR horizon is 10 days. We perform a Delta-Gamma approximation via Monte Carlo simulation[13]. To this purpose, we notice that the Cholesky decomposition of the covariance matrix is given in Table 28. The simulated per-

Cholesky Decomposition		
0.0291117	0	0
0.0204735	0.030261	0
0.0255482	0.012533	0.011962

Table 28: Cholesky Decomposition of the covariance matrix in Table 27.

centage price changes in the first simulation are given in the third column of Table 29. The first column refers to a vector of independent Gaussian random variables with mean 0 and variance equal to the VaR horizon. The second column refers to the simulated returns on the three commodities. It has been obtained by multiplying the Cholesky decomposition by the vector of independent Gaussian random variables.

The simulated portfolio P&L according to the Delta-Gamma approximation is

---

[13]Monte Carlo simulation in the multivariate Gaussian case is presented in Appendix.

$\mathcal{N}\left(0, \frac{10}{250}\right)$	$\mathbf{r}\left(t, t + \frac{10}{250}\right)$	$d\mathbf{P}/\mathbf{P}$
-0.16019	-0.00466	-0.00465
-0.44117	-0.01663	-0.01649
-0.03469	-0.01004	-0.00999

Table 29: Example of simulation of percentage price changes on the three given commodities. First column: vector of simulated independent Gaussian random variables with zero mean and variance equal to 10/250. Second column: simulated returns obtained by the product between the Cholesky matrix in Table and the first column of the present Table. Third column: simulated percentage price changes.

therefore obtained as

$$
= -190.11357 \times \frac{10}{250}
$$

$$
+ \begin{bmatrix} 101.48743 & -103.90143 & 96.11865 \end{bmatrix} \begin{bmatrix} -0.465\% \\ -1.649\% \\ -0.999\% \end{bmatrix}
$$

$$
+ \frac{1}{2} \begin{bmatrix} -0.465\% & -1.649\% & -0.999\% \end{bmatrix} \begin{bmatrix} 237.10 & 0 & 0 \\ 0 & 1415.45 & 0 \\ 0 & 0 & 474.31 \end{bmatrix} \begin{bmatrix} -0.465\% \\ -1.649\% \\ -0.999\% \end{bmatrix}
$$

$$
= -7.233.
$$

If we repeat the simulations 100,000 times in all, we can compute the empirical quantile of the simulated P&L. In our example, the 99% VaR turns out to be 191.89$. The simulated P&L distribution arising from the Delta-Gamma approximation is illustrated in Figure 20.

7.3 Issues with the Delta-Gamma approximation

The main issues related to the Taylor's approximation are related to the accuracy of the Taylor's formula and to the computation of the Greeks. On the other hand, one of the main advantages of this approximation is that it can be used independently of the assumed model for the risk-drivers and it requires very little computational resources.

In general, the Delta-Gamma approximation is valid only for fairly small changes. This is shown in Figure 21 where we plot the Black-Scholes option price as function of the underlying asset price and its approximation to the first and second order. Due to the convexity of the option price, here the linear approximation overestimates the variation in the option
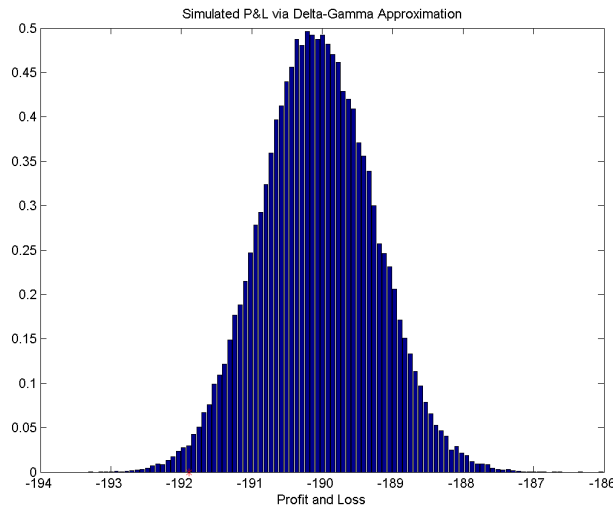
Figure 20: Simulated P&L distribution in the Delta-Gamma approximation of Example 44. The red asterisk gives us the 99% VaR level.

| Option VaR ($\alpha = 95\%$) and Time Horizon (in days)						
Days	1	5	10	20	50	82.5
Linear	4.98	11.21	16.20	22.88	38.06	49.72
Quadratic	4.71	9.94	13.66	18.18	26.66	32.05
Full	4.70	9.83	13.34	17.25	22.09	22.47
Exact	4.66	9.83	13.25	17.35	22.07	22.47

Table 30: 95%-VaR for a call option for diffent time horizons and under different methods: linear approximation, quadratic approximation, full revaluation via Monte Carlo (10,000 simulations) and exact VaR number. Figures are in dollar terms.

P&L. This is confirmed by looking at the the option VaR in Table 30, varying the time horizon. Further, in the case of short maturity at-the-money options and options with discontinuous payoffs like digitals and barriers, the Greeks can change abruptly near the discontinuity points. In these cases, the second-order approximation can be often unsatisfactory and the full revaluation is the only reliable alternative. Important contributions on how to improve and speed up the delta-gamma approximations include Cardenas et al. [12] and Rouvinez [24].

Concerning the computation of the Greeks, for plain vanilla products, there exist analytical pricing function so that the Greeks can be also computed in analytical way. If the analytical computation of the Greeks is too
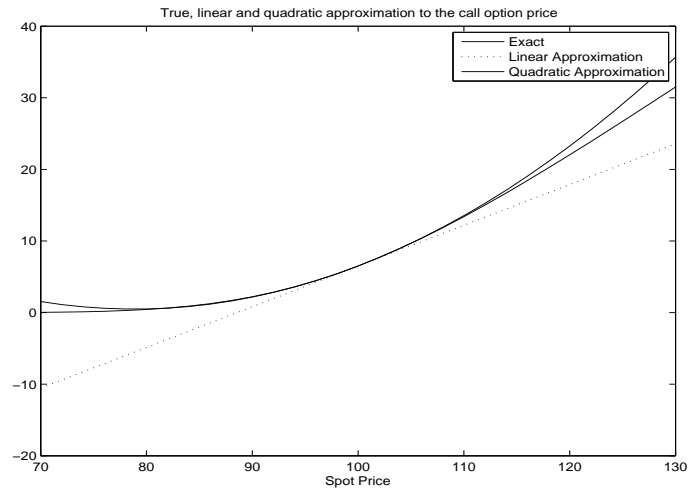
Figure 21: Relation between true option price (using the Black-Scholes formula), linear and quadratic approximation.
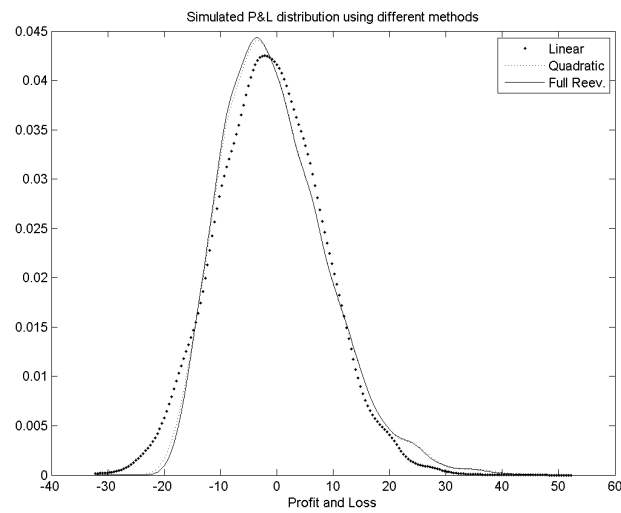
Figure 22: Density of the option P&L using the exact density, linear and quadratic approximation. (10 days horizon).

79

complicated, a possible alternative is to estimate the option Theta, Delta and Gamma numerically, using finite difference approximations as follows

$$\Theta \simeq \frac{C(P, t + dt) - C(P, t)}{dt},$$

$$\Delta_i \simeq \frac{C(P_i + dP_i, t) - C(P_i - dP_i, t)}{2dP_i},$$

$$\Gamma_{ii} \simeq \frac{C(P_i + dP_i, t) - 2C(P_i, t) + C(P_i - dP_i, t)}{(dP_i)^2}.$$

Example 45 (Computing Greeks by Finite Difference) *Let us consider the call option of Example 40.*

- *The Theta can be estimated by shortening the time to maturity by a small amount, 10^{-5} say*

$$\frac{C(P, t + 10^{-5}) - C(P, t)}{10^{-5}} = \frac{22.467728 - 22.46805}{10^{-5}} = -32.463,$$

which is comparable to the true Theta of -32.4625.

- *The Delta and Gamma can be computed by shifting up and down the current level price by a small amount, 10^{-5} say, so that*

$$C(P + dP, t) = 22.468059, \ C(P - dP, t) = 22.468047.$$

- *Therefore*

$$\Delta = \frac{22.468059 - 22.468047}{2 \times 10^{-5}} = 0.612577,$$

$$\Gamma = \frac{22.468059 - 2 \times 22.468053 + 22.468047}{(10^{-5})^2} = 0.009094,$$

which can be compared to the exact values 0.612577 and 0.00857, respectively.

The computation of the Greeks is more problematic when the pricing function is not analytical, but it requires approximation via numerical method, like Monte Carlo simulation. In this case, the use of finite difference can return quite inaccurate results so that the use of more advanced procedures is required. Suitable techniques include the pathwise method and likelihood ratio method, both of which are reviewed in Chapter 7 of Glasserman [19]. The recent adjoint method seems well suited to applications requiring sensitivities to a large number of parameters, see Giles and Glasserman [20] and Capriotti and Giles [11].

References

[1] Alexander, C. (2008) Market Risk Analysis, Volume IV: Value at Risk Models. Wiley.

[2] Artzner, P., F. Delbaen, J.-M. Eber and D. Heath (1997) Thinking coherently. Risk 10 (November): 68-71.

[3] Artzner, P., F. Delbaen, J.-M. Eber and D. Heath (1999) Coherent measures of risk. Mathematical Finance 9 (November): 203-228.

[4] Ballotta, L., Fusai, G. and Marena, M. (2016). A Gentle Introduction to Default Risk and Counterparty Credit Modelling. Available at SSRN: https://ssrn.com/abstract=2816355.

[5] Berkowitz, J. (2001) Testing Density Forecasts With Applications to Risk Management. Journal of Business and Economic Statistics, 19, 465-474.

[6] Berkowitz, J. (2002) Testing Assumptions. Risk 15, May, 77-80.

[7] Britten-Jones, M. and Schaefer S. M. (1999). Non-Linear Value-at-Risk, European Finance Review 2: 161-187, 1999.

[8] Caldana R. and Fusai G. (2013). A general closed-form spread option pricing formula, Journal of Banking & Finance, 37, 4893-4906.

[9] Caldana, R., Fusai, G., Gnoatto, A., and Grasselli, M. (2016) General Closed-Form Basket Option Pricing Bounds. Quantitative Finance Volume 16, Issue 4, 535-554. Available at SSRN: http://ssrn.com/abstract=2376134 or http://dx.doi.org/10.2139/ssrn.2376134.

[10] Campbell, S. D. (2006) A review of backtesting and backtesting procedures, Journal of Risk, 9 (2), 1-17.

[11] Capriotti, L. and M. Giles (2010) Fast Correlation Greeks by Adjoint Algorithmic Differentiation, Risk Magazine, April.

[12] Cardenas, J., E. Fruchard, J.-F. Picron, C. Reyes, K. Walters and W. Yang (1999) Monte Carlo within a day. Risk Magazine 12 (February): 55-59.

[13] Christoffersen P. (2011). Elements of Financial Risk Management, Academic Press, 2 edition.

[14] Crnkovic C., and Drachman J., (1996) Quality Control, Risk 9, September, 139-143. Reprinted in VaR: Understanding and Applying Value-at-Risk, 1997, London, Risk Publications.

[15] Efron, B. (1979) Bootstrap Methods: Another Look at the Jackknife, The Annals of Statistics 7, 1-26.

[16] Diebold, F. X., Gunther, T. A., and Tay, A. S. (1998) Evaluating density forecasts with applications to financial risk management, International Economic Review 39, 863-883.

[17] Dowd, K. (2002). An Introduction to Market Risk Measurement, Wiley Finance.

[18] Efimova, O. and Serletis, A. (2014). Energy markets volatility modelling using GARCH, Energy Economics 43, 264-273.

[19] Glasserman P. (2003) Monte Carlo Methods in Financial Engineering Springer-Verlag, New York.

[20] Giles M., and Glasserman, P., (2006) Smoking adjoints: fast Monte Carlo Greeks, Risk Magazine 19 (January), 92-96.

[21] Jarque, C. M. and Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals, Economics Letters 6 (3), 255-259.

[22] Jorion, P. (2007) Value at Risk, Mc Graw Hill, Singapore, 3rd edition.

[23] Politis, D. and J. P. Romano (1994) The Stationary Bootstrap, Journal of the American Statistical Association 89, 1303-1313.

[24] Rouvinez, C. (1997) Going Greek with VaR. Risk 10 (February): 57-65.

[25] Stultz, R. (1996) Rethinking Risk Management, Journal of Applied Corporate Finance, Volume 9, Issue 3, pages 8-25, Fall 1996.

A   Monte Carlo simulation of Gaussian random variables

There are several methods to simulate a standard Gaussian random variable Z. A very well known method is based on the inverse CDF of the standard Gaussian distribution and on the generation of a Uniform random variable, as follows

- Simulate a Uniform U;

- Use the inverse cumulative density function $\Phi_{0,1}(x)$ of the standard normal random
$$Z = \Phi_{0,1}^{-1}(U).$$

The procedure is illustrated in Figure 23 and can be easily implemented in Excel by inputing in a cell the formula

```
=NORMSINV(RAND())
```

If we need to simulate a Gaussian random variable with mean μ and standard deviation σ, we apply to the simulated standard Gaussian random variable the linear transformation

$$\mu + \sigma Z.$$

In the multivariate case if we need to simulate a random vector

$$\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}),$$

we proceed as follows

1. Given the covariance matrix Σ, we compute its Cholesky decomposition, i.e. a lower triangular matrix $\mathbf{A}$ such that

$$\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}'.$$

2. We simulate a vector of standard indepedent Gaussian random variables $\mathbf{Z}$. The components of this vector can be obtained by simulating N standard Gaussian random variables.

3. We simulate from the desired multivariate Gaussian random variable by setting

$$\mathbf{X} = \boldsymbol{\mu} + \mathbf{A}\mathbf{Z}.$$
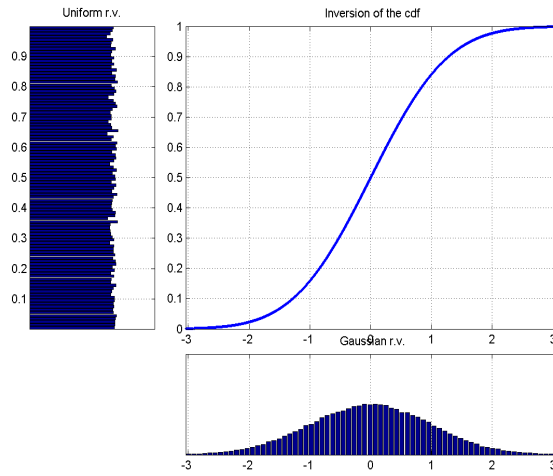
Figure 23: Scheme of the simulation of a standard Gaussian random variable via the inverse method.

Example 46 (Cholesky decomposition of a 2x2 matrix) *Let*

$$\Sigma = \begin{bmatrix} 0.04 & 0.024 \\ 0.024 & 0.09 \end{bmatrix}.$$

In order to find its Cholesky decomposition, we need to look for a 2x2 lower triangular matrix **A** *such that*

$$\begin{bmatrix} 0.04 & 0.024 \\ 0.024 & 0.09 \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} \\ 0 & a_{22} \end{bmatrix} = \begin{bmatrix} a_{11}^2 & a_{11}a_{21} \\ a_{11}a_{21} & a_{21}^2 + a_{22}^2 \end{bmatrix},$$

i.e. we have to set

$$\begin{cases} a_{11}^2 = 0.04 \\ a_{11}a_{21} = 0.024 \\ a_{21}^2 + a_{22}^2 = 0.09 \end{cases} \implies \begin{cases} a_{11} = \sqrt{0.04} = 0.2 \\ a_{21} = \frac{0.024}{0.2} = 0.12 \\ a_{22} = \sqrt{0.09 - (0.12)^2} = \sqrt{0.0756} = 0.27495 \end{cases}$$

and therefore

$$\mathbf{A} = \begin{bmatrix} 0.2 & 0 \\ 0.12 & 0.2749545 \end{bmatrix}.$$

It can be verified that if

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix},$$

then

$$\mathbf{A} = \begin{bmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2\sqrt{1-\rho^2} \end{bmatrix}.$$

This means that in order to perform Monte Carlo simulation, we need to simulate Z_1 and Z_2 independently and according to a standard normal random variable, and then to set (if the means are not zero)

$$X_1 = \mu_1 + \sigma_1 Z_1,$$

$$X_2 = \mu_2 + \rho\sigma_2 Z_1 + \sigma_2\sqrt{1-\rho^2}Z_2.$$

In our numerical example, correlated Gaussian r.v's are simulated according to

$$X_1 = \mu_1 + 0.2 \times Z_1,$$

$$X_2 = \mu_2 + 0.4 \times 0.3 \times Z_1 + 0.3 \times \sqrt{1-0.4^2} \times Z_2.$$

If $\mu_1 = 1, \mu_2 = 2$, and the simulated standard unit Gaussian random numbers are $Z_1 = 0.5376, Z_2 = 1.8338$, the simulated correlated Gaussian random numbers are

$$X_1 = 1 + 0.2 \times 0.5376 = 2.0752,$$

$$X_2 = 2 + 0.4 \times 0.3 \times 0.5376 + 0.3 \times \sqrt{1-0.4^2} \times 1.8338 = 2.5687.$$

In the case of a $N \times N$ covariance matrix, the Cholesky decomposition has entries

$$A_{jj} = \sqrt{\sigma_j^2 - \sum_{k=1}^{j-1} \sigma_{jk}^2},$$

$$A_{ij} = \frac{1}{A_{jj}}\left(\sigma_{ij} - \sum_{k=1}^{j-1}\sigma_{ik}\sigma_{jk}\right), \text{ for } i > k.$$

The expression under the square root is always positive if Σ is real and positive definite.