2024

FRM®

EXAM PART II

Market Risk Measurement and Management

GARP®

FRM | Financial Risk Manager

GARP®

FRM® | Financial Risk Manager

2024

FRM®

EXAM PART II

Market Risk Measurement and Management

Pearson

Excerpts taken from:

Options, Futures, and Other Derivatives, Tenth Edition by John C. Hull
Copyright © 2017, 2015, 2012, 2009, 2006, 2003, 2000 by Pearson Education, Inc.
New York, New York 10013

Grateful acknowledgment is made to the following sources for permission to reprint material copyrighted or controlled by them:

"Estimating Market Risk Measures," "Non-Parametric Approaches," and "Parametric Approaches (III): Extreme Value" by Kevin Dowd, reprinted from *Measuring Market Risk*, Second Edition (2005), by permission of John Wiley & Sons, Inc. All rights reserved. Used under license from John Wiley & Sons, Inc.

"Back Testing VaR" and "VaR Mapping," by Philippe Jorion, reprinted from *Value at Risk: The New Benchmark for Managing Financial Risk*, Third Edition (2007), by permission of The McGraw Hill Companies.

"Messages from the Academic Literature on Risk Measurement for the Trading Book," Working Paper No. 19 January 2011, reprinted by permission of the Basel Committee on Banking Supervision.

"Some Correlation Basics: Definitions, Applications, and Terminology," "Empirical Properties of Correlation: How Do Correlations Behave in the Real World?," and "Financial Correlation Modeling—Bottom Up Approaches," by Gunter Meissner, reprinted from *Correlation Risk Modeling and Management*, Second Edition (2019), by permission of Risk Books/InfoPro Digital Services, Ltd.

"Empirical Approaches to Risk Metrics and Hedges," "The Science of Term Structure Models," "The Evolution of Short Rates and the Shape of the Term Structure," "The Art of Term Structure Models: Drift," and "The Art of Term Structure Models: Volatility and Distribution," by Bruce Tuckman and Angel Serrat, reprinted from *Fixed Income Securities: Tools for Today's Markets*, Third Edition (2012), by permission of John Wiley & Sons, Inc. All rights reserved. Used under license from John Wiley & Sons, Inc.

"Fundamental Review of the Trading Book," by John C. Hull, reprinted from *Risk Management and Financial Institutions*, Fifth Edition (2018), by permission of John Wiley & Sons, Inc. All rights reserved. Used under license from John Wiley & Sons, Inc.

Learning Objectives provided by the Global Association of Risk Professionals.

Pearson Education, Inc., 221 River Street, Hoboken, NJ 07030
A Pearson Education Company
www.pearsoned.com

Printed in the United States of America

ScoutAutomatedPrintCode

00033038-00000005 / A103001321011
EEB/SK

Pearson

ISBN 10: 0-13-829218-3
ISBN 13: 978-0-13-829218-8

Contents

PREFACE

I want to thank you on behalf of GARP's Board of Trustees and our professional certification program staff for your support of the Financial Risk Manager (FRM®) program.

It's gratifying to see that in the 26 years since the first FRM examination, the FRM program has become the global standard for educating and credentialing financial risk management professionals. Its worldwide effects in furthering the understanding and acceptance of financial risk management have been highly positive and, in many ways, transformative.

COVID is thankfully in the rearview mirror. We now can be much more flexible in expanding—and in certain instances re-focusing and updating—the FRM program to address the many new challenges encountered by financial institutions globally.

Our FRM program advisory committee, consisting of senior risk professionals from around the world, that meets regularly to debate and settle the FRM program's subject coverage, has found no shortage of subjects for inclusion in the FRM curriculum.

One of the advisory committee's more-material challenges is to understand and assess where the global financial services industry is headed, and then identify issues and subjects most important for risk management professionals.

The FRM advisory committee also recommends how the FRM program covers subject matter. Its objective is to ensure that candidates who complete the FRM program successfully can be confident that their skills have been assessed objectively, and that they possess the requisite knowledge to succeed as a risk management professional anywhere in the world.

The FRM program's coverage is dynamic. The advisory committee reacts to and tries to anticipate market changes, global economic trends, technological advances, and regulatory adjustments; and assesses how these will affect the necessary knowledge and skill sets of a risk management professional.

The biggest change to the program's coverage for 2024 revolves around credit risk measurement and management. About two-thirds of the subject readings in *Credit Risk Measurement and Management* were updated for 2024.

Notably in 2023, GARP expanded the FRM program's coverage of operational resilience, an issue of rapidly growing importance around the world. Materials deal with structural vulnerabilities and areas of the financial system that may be under stress. The transmission of shocks to the financial system, and the assessment, modeling, and measurement of potential points of failure are other important covered concepts.

Also notable in 2023, GARP added two chapters on machine learning (ML) in the FRM Part I *Quantitative Analysis* book. These chapters not only introduce the ML methods risk managers need to understand, but also address key issues associated with artificial intelligence (AI) and ML, including transparency, interpretability, and explainability; data considerations; and risks that arise from the use of AI/ML, including the potential for bias, discrimination, and unethical behavior.

Throughout the FRM curriculum, GARP aims, wherever possible, to present lessons learned from noteworthy current events to contextualize program content and give FRM candidates critical insight.

As you will see from reviewing the program's coverage and readings, it keeps up with a world that is becoming more interconnected and complex by the day.

GARP is committed to offering a program that is dynamic, sophisticated, and responsive to the needs of financial institutions and risk professionals around the world.

We wish you the very best as you study for the FRM exams. And much success in your career as a risk-management professional.

Yours truly,

Richard Apostolik

President & CEO

FRM® COMMITTEE

Chairperson

Nick Strange, FCA
Senior Technical Advisor, Operational Risk & Resilience,
Prudential Regulation Authority, Bank of England

Members

Richard Apostolik
President and CEO, GARP

Richard Brandt
MD, Operational Risk Management, Citigroup

Julian Chen, FRM
SVP, FRM Program Manager, GARP

Chris Donohue, PhD
MD, GARP Benchmarking Initiative, GARP

Donald Edgar, FRM
MD, Risk & Quantitative Analysis, BlackRock

Hervé Geny
Former Group Head of Internal Audit, London Stock Exchange Group

Apama Gupta
Professor of Quantitative Finance
Associate Dean, Academic Affairs
A.W. Lawrence Professional Excellence Fellow
Co-Director and Site Director, NSF IUCRC CRAFT
Lally School of Management
Rensselaer Polytechnic Institute

John Hull
Senior Advisor
Maple Financial Professor of Derivatives and Risk Management,
Joseph L. Rotman School of Management, University of Toronto

Keith Isaac, FRM
VP, Capital Markets Risk Management, TD Bank Group

William May
SVP, Global Head of Certifications and Educational Programs, GARP

Attilio Meucci, PhD, CFA
Founder, ARPM

Victor Ng, PhD
Chairman, Audit and Risk Committee
Former MD, Head of Risk Architecture, Goldman Sachs

Matthew Pritsker, PhD
Senior Financial Economist and Policy Advisor/Supervision, Regulation, and Credit, Federal Reserve Bank of Boston

Samantha C. Roberts, PhD, FRM, SCR
Instructor and Consultant, Risk Modeling and Analytics

Til Schuermann, PhD
Partner, Oliver Wyman

Evan Sekeris, PhD
Head of Non-Financial Risk, MUFG

Sverrir Þorvaldsson, PhD, FRM
Senior Quant, SEB

Estimating Market Risk Measures

An Introduction and Overview

1

Learning Objectives

After completing this reading, you should be able to:

- Estimate VaR using a historical simulation approach.
- Estimate VaR using a parametric approach for both normal and lognormal return distributions.
- Estimate the expected shortfall given profit and loss (P&L) or return data.
- Estimate risk measures by estimating quantiles.
- Evaluate estimators of risk measures by estimating their standard errors.
- Interpret quantile-quantile (QQ) plots to identify the characteristics of a distribution.

Excerpt is Chapter 3 of Measuring Market Risk, *Second Edition, by Kevin Dowd.*

This chapter provides a brief introduction and overview of the main issues in market risk measurement. Our main concerns are:

- *Preliminary data issues*: How to deal with data in profit/loss form, rate-of-return form, and so on.
- *Basic methods of VaR estimation*: How to estimate simple VaRs, and how VaR estimation depends on assumptions about data distributions.
- How to estimate coherent risk measures.
- How to gauge the precision of our risk measure estimators by estimating their standard errors.
- *Overview*: An overview of the different approaches to market risk measurement, and of how they fit together.

We begin with the data issues.

1.1 DATA

Profit/Loss Data

Our data can come in various forms. Perhaps the simplest is in terms of profit/loss (or P/L). The P/L generated by an asset (or portfolio) over the period t, P/L_t, can be defined as the value of the asset (or portfolio) at the end of t plus any interim payments D_t minus the asset value at the end of $t-1$:

$$P/L_t = P_t + D_t - P_{t-1} \tag{1.1}$$

If data are in P/L form, positive values indicate profits and negative values indicate losses.

If we wish to be strictly correct, we should evaluate all payments from the same point of time (i.e., we should take account of the time value of money). We can do so in one of two ways. The first way is to take the present value of P/L_t evaluated at the end of the previous period, $t-1$:

$$\text{Present Value } (P/L_t) = \frac{(P_t + D_t)}{(1+d)} - P_{t-1} \tag{1.2}$$

where d is the discount rate and we assume for convenience that D_t is paid at the end of t. The alternative is to take the forward value of P/L_t evaluated at the end of period t:

$$\text{Forward Value } (P/L_t) = P_t + D_t - (1+d)P_{t-1} \tag{1.3}$$

which involves compounding P_{t-1} by d. The differences between these values depend on the discount rate d, and will be small if the periods themselves are short. We will ignore these differences to simplify the discussion, but they can make a difference in practice when dealing with longer periods.

Loss/Profit Data

When estimating VaR and ES, it is sometimes more convenient to deal with data in loss/profit (L/P) form. L/P data are a simple transformation of P/L data:

$$L/P_t = -P/L_t \tag{1.4}$$

L/P observations assign a positive value to losses and a negative value to profits, and we will call these L/P data 'losses' for short. Dealing with losses is sometimes a little more convenient for risk measurement purposes because the risk measures are themselves denominated in loss terms.

Arithmetic Return Data

Data can also come in the form of arithmetic (or simple) returns. The arithmetic return r_t is defined as:

$$r_t = \frac{P_t + D_t - P_{t-1}}{P_{t-1}} = \frac{P_t + D_t}{P_{t-1}} - 1 \tag{1.5}$$

which is the same as the P/L over period t divided by the value of the asset at the end of $t-1$.

In using arithmetic returns, we implicitly assume that the interim payment D_t does not earn any return of its own. However, this assumption will seldom be appropriate over long periods because interim income is usually reinvested. Hence, arithmetic returns should not be used when we are concerned with long horizons.

Geometric Return Data

Returns can also be expressed in geometric (or compound) form. The geometric return R_t is

$$R_t = \ln\left(\frac{P_t + D_t}{P_{t-1}}\right) \tag{1.6}$$

The geometric return implicitly assumes that interim payments are continuously reinvested. The geometric return is often more economically meaningful than the arithmetic return, because it ensures that the asset price (or portfolio value) can never become negative regardless of how negative the returns might be. With arithmetic returns, on the other hand, a very low realized return—or a high loss—implies that the asset value P_t can become negative, and a negative asset price seldom makes economic sense.[1]

The geometric return is also more convenient. For example, if we are dealing with foreign currency positions, geometric returns will give us results that are independent of the reference

[1] This is mainly a point of principle rather than practice. In practice, any distribution we fit to returns is only likely to be an approximation, and many distributions are ill-suited to extreme returns anyway.

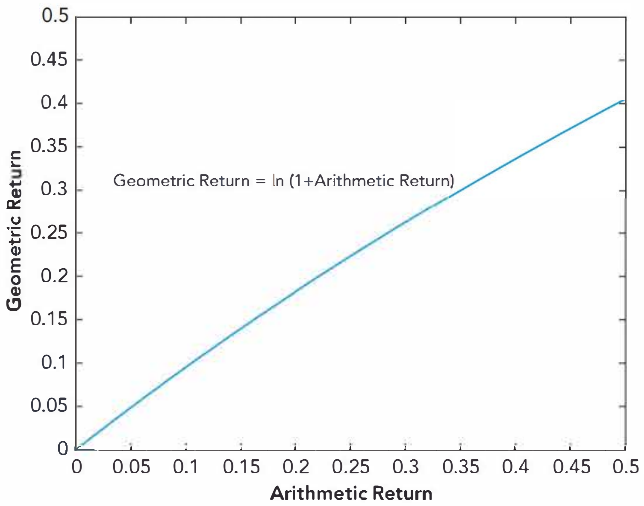

Figure 1.1 **Geometric and arithmetic returns.**

currency. Similarly, if we are dealing with multiple periods, the geometric return over those periods is the sum of the one-period geometric returns. Arithmetic returns have neither of these convenient properties.

The relationship of the two types of return can be seen by rewriting Equation (1.6) (using a Taylor's series expansion for the natural log) as:

$$R_t = \ln\left(\frac{P_t + D_t}{P_{t-1}}\right) = \ln(1 + r_t) = r_t - \frac{1}{2}r_t^2 + \frac{1}{3}r_t^3 - \cdots \quad (1.7)$$

from which we can see that $R_t \approx r_t$ provided that returns are 'small'. This conclusion is illustrated by Figure 1.1, which plots the geometric return R_t against its arithmetic counterpart r_t. The difference between the two returns is negligible when both returns are small, but the difference grows as the returns get bigger—which is to be expected, as the geometric return is a log function of the arithmetic return. Since we would expect returns to be low over short periods and higher over longer periods, the difference between the two types of return is negligible over short periods but potentially substantial over longer ones. And since the geometric return takes account of earnings on interim income, and the arithmetic return does not, we should always use the geometric return if we are dealing with returns over longer periods.

Example 1.1 Arithmetic and Geometric Returns

If arithmetic returns r_t over some period are 0.05, Equation (1.7) tells us that the corresponding geometric returns are $R_t = \ln(1 + r_t) = \ln(1.05) = 0.0488$. Similarly, if geometric returns R_t are 0.05, Equation (1.7) implies that arithmetic returns are $1 + r_t = \exp(R_t) \Rightarrow r_t = \exp(R_t) - 1 = \exp(0.05) - 1 = 0.0513$. In both cases the arithmetic return is close to, but a little higher than, the geometric return—and this makes intuitive sense when one considers that the geometric return compounds at a faster rate.

1.2 ESTIMATING HISTORICAL SIMULATION VaR

The simplest way to estimate VaR is by means of historical simulation (HS). The HS approach estimates VaR by means of ordered loss observations.

Suppose we have 1000 loss observations and are interested in the VaR at the 95% confidence level. Since the confidence level implies a 5% tail, we know that there are 50 observations in the tail, and we can take the VaR to be the 51st highest loss observation.[2]

We can estimate the VaR on a spreadsheet by ordering our data and reading off the 51st largest observation from the spreadsheet. We can also estimate it more directly by using the 'Large' command in Excel, which gives us the *k*th largest value in an array. Thus, if our data are an array called 'Loss_data', then our VaR is given by the Excel command 'Large(Loss_data,51)'. If we are using MATLAB, we first order the loss/profit data using the 'Sort()' command (i.e., by typing 'Loss_data = Sort(Loss_data)'); and then derive the VaR by typing in 'Loss_data(51)' at the command line.

More generally, if we have *n* observations, and our confidence level is α, we would want the $(1 - \alpha).n + 1$th highest observation, and we would use the commands 'Large(Loss_data,(1 − *alpha*)**n* + 1)' using Excel, or 'Loss_data((1 − *alpha*)**n* + 1)' using MATLAB, provided in the latter case that our 'Loss_data' array is already sorted into ordered observations.[3]

[2] In theory, the VaR is the quantile that demarcates the tail region from the non-tail region, where the size of the tail is determined by the confidence level, but with finite samples there is a certain level of arbitrariness in how the ordered observations relate to the VaR itself—that is, do we take the VaR to be the 50th observation, the 51st observation, or some combination of them? However, this is just an issue of approximation, and taking the VaR to be the 51st highest observation is not unreasonable.

[3] We can also estimate HS VaR using percentile functions such as the 'Percentile' function in Excel or the 'prctile' function in MATLAB. However, such functions are less transparent (i.e., it is not obvious to the reader how the percentiles are calculated), and the Excel percentile function can be unreliable.

An example of an HS VaR is given in Figure 1.2. This figure shows the histogram of 1000 hypothetical loss observations and the 95%VaR. The figure is generated using the 'hsvarfigure' command in the MMR Toolbox. The VaR is 1.704 and separates the top 5% from the bottom 95% of loss observations.

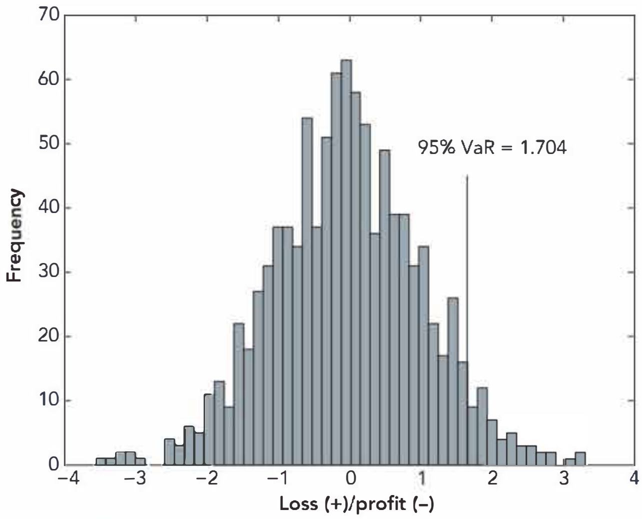

Figure 1.2 **Historical simulation VaR.**

Note: Based on 1000 random numbers drawn from a standard normal L/P distribution, and estimated with 'hsvarfigure' function.

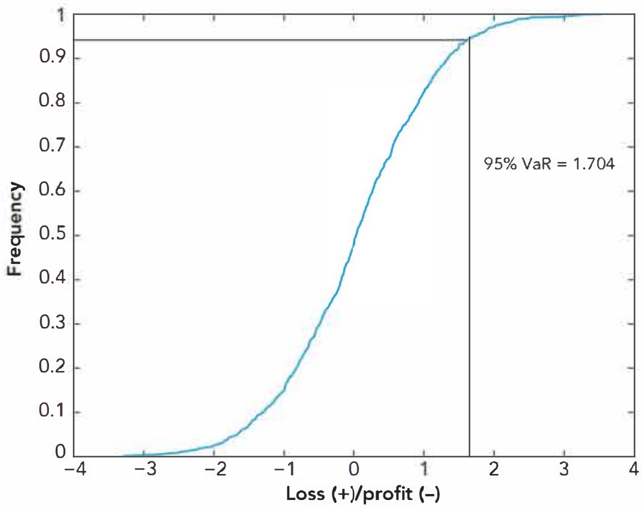

Figure 1.3 **Historical simulation via an empirical cumulative frequency function.**

Note: Based on the same data as Figure 1.2.

In practice, it is often helpful to obtain HS VaR estimates from a cumulative histogram, or empirical cumulative frequency function. This is a plot of the ordered loss observations against their empirical cumulative frequency (e.g., so if there are n observations in total, the empirical cumulative frequency of the ith such ordered observation is i/n). The empirical cumulative frequency function of our earlier data set is shown in Figure 1.3. The empirical frequency function makes it very easy to obtain the VaR: we simply move up the cumulative frequency axis to where the cumulative frequency equals our confidence level, draw a horizontal line along to the curve, and then draw a vertical line down to the x-axis, which gives us our VaR.

1.3 ESTIMATING PARAMETRIC VAR

We can also estimate VaR using parametric approaches, the distinguishing feature of which is that they require us to explicitly specify the statistical distribution from which our data observations are drawn. We can also think of parametric approaches as fitting curves through the data and then reading off the VaR from the fitted curve.

In making use of a parametric approach, we therefore need to take account of both the statistical distribution and the type of data to which it applies.

Estimating VaR with Normally Distributed Profits/Losses

Suppose that we wish to estimate VaR under the assumption that P/L is normally distributed. In this case our VaR at the confidence level α is:

$$\alpha VaR = -\mu_{P/L} + \sigma_{P/L} z_\alpha \quad (1.8)$$

where z_α is the standard normal variate corresponding to α, and $\mu_{P/L}$ and $\sigma_{P/L}$ are the mean and standard deviation of P/L. Thus, z_α is the value of the standard normal variate such that α of the probability density mass lies to its left, and $1 - \alpha$ of the probability density mass lies to its right. For example, if our confidence level is 95%, $z_\alpha = z_{0.95}$ will be 1.645.

In practice, $\mu_{P/L}$ and $\sigma_{P/L}$ would be unknown, and we would have to estimate VaR based on estimates of these parameters. Our VaR estimate, αVaR^e, would then be:

$$\alpha VaR^e = -m_{P/L} + s_{P/L} z_\alpha \quad (1.9)$$

where $m_{P/L}$ and $s_{P/L}$ are estimates of the mean and standard deviation of P/L.

Figure 1.4 shows the 95% VaR for a normally distributed P/L with mean 0 and standard deviation 1. Since the data are in P/L form, the VaR is indicated by the negative of the cut off point between the lower 5% and the upper 95% of P/L observations. The actual VaR is the negative of −1.645, and is therefore 1.645.

If we are working with normally distributed L/P data, then $\mu_{L/P} = -\mu_{P/L}$ and $\sigma_{L/P} = \sigma_{P/L}$, and it immediately follows that:

$$\alpha VaR = \mu_{L/P} + \sigma_{L/P} z_\alpha \quad (1.10a)$$

$$\alpha VaR^e = m_{L/P} + s_{L/P} z_\alpha \quad (1.10b)$$

Figure 1.5 illustrates the corresponding VaR. This figure gives the same information as Figure 1.4, but is a little more straightforward to interpret because the VaR is defined in units of losses (or 'lost money') rather than P/L. In this case, the VaR is given by the point on the *x*-axis that cuts off the top 5% of the pdf mass from the bottom 95% of pdf mass. If we prefer to work with the cumulative density function, the VaR is the *x*-value that corresponds to a cdf value of 95%. Either way, the VaR is again 1.645, as we would (hopefully) expect.

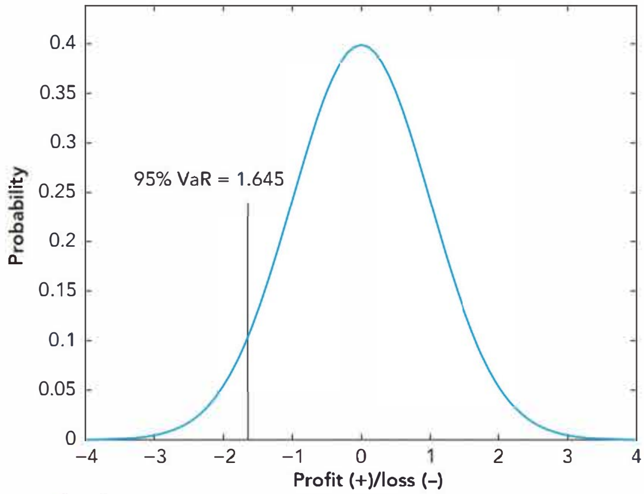

Figure 1.4 VaR with standard normally distributed profit/loss data.

Note: Obtained from Equation (1.9) with $\mu_{P/L} = 0$ and $\sigma_{P/L} = 1$. Estimated with the 'normalvarfigure' function.

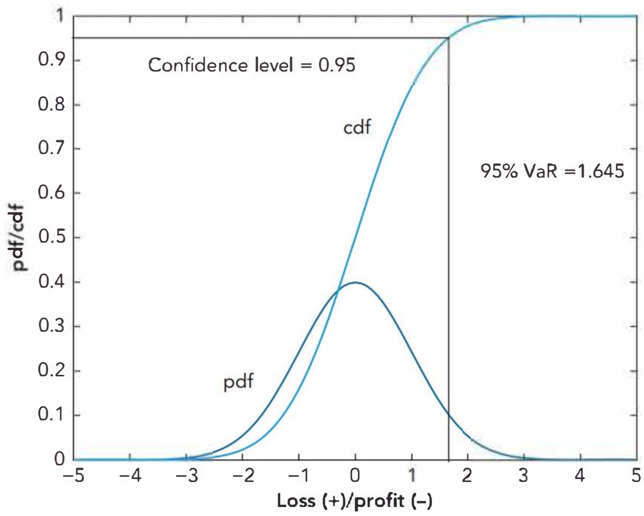

Figure 1.5 VaR with normally distributed loss/profit data.

Note: Obtained from Equation (1.10a) with $\mu_{L/P} = 0$ and $\sigma_{L/P} = 1$.

Example 1.2 VaR with Normal P/L

If P/L over some period is normally distributed with mean 10 and standard deviation 20, then (by Equation (1.8)) the 95% VaR is $-10 + 20z_{0.95} = -10 + 20 \times 1.645 = 22.9$. The corresponding 99% VaR is $-10 + 20z_{0.99} = -10 + 20 \times 2.326 = 36.52$.

Estimating VaR with Normally Distributed Arithmetic Returns

We can also estimate VaR making assumptions about returns rather than P/L. Suppose then that we assume that arithmetic returns are normally distributed with mean μ_r and standard deviation σ_r. To derive the VaR, we begin by obtaining the critical value of r_t, r^*, such that the probability that r_t exceeds r^* is equal to our confidence level α. r^* is therefore:

$$r^* = \mu_r - \sigma_r z_\alpha \quad (1.11)$$

Since the actual return r_t is the loss/profit divided by the earlier asset value, P_{t-1}, it follows that:

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}} = -\frac{Loss_t}{P_{t-1}} \quad (1.12)$$

Substituting r^* for r_t then gives us the relationship between r^* and the VaR:

$$r_t^* = \frac{P_t^* - P_{t-1}}{P_{t-1}} = \frac{VaR}{P_{t-1}} \tag{1.13}$$

Substituting Equation (1.11) into Equation (1.13) and rearranging then gives us the VaR itself:

$$\alpha VaR = -(\mu_r - \sigma_r z_\alpha)P_{t-1} \tag{1.14}$$

Equation (1.14) will give us equivalent answers to our earlier VaR equations. For example, if we set $\alpha = 0.95$, $\mu_r = 0$, $\sigma_r = 1$ and $P_{t-1} = 1$, which correspond to our earlier illustrative P/L and L/P parameter assumptions, αVaR is 1.645: the three approaches give the same results, because all three sets of underlying assumptions are equivalent.

Example 1.3 VaR with Normally Distributed Arithmetic Returns

Suppose arithmetic returns r_t over some period are distributed as normal with mean 0.1 and standard deviation 0.25, and we have a portfolio currently worth 1. Then (by Equation (1.14)) the 95% VaR is $-0.1 + 0.25 \times 1.645 = 0.331$, and the 99% VaR is $-0.1 + 0.25 \times 2.326 = 0.482$.

Estimating Lognormal VaR

Each of the previous approaches assigns a positive probability of the asset value, P_t, becoming negative, but we can avoid this drawback by working with geometric returns. Now assume that geometric returns are normally distributed with mean μ_R and standard deviation σ_R. If D_t is zero or reinvested continually in the asset itself (e.g., as with profits reinvested in a mutual fund), this assumption implies that the natural logarithm of P_t is normally distributed, or that P_t itself is lognormally distributed. The lognormal distribution is explained in Box 1.1, and a lognormal asset price is shown in Figure 1.6: observe that the price is always non-negative, and its distribution is skewed with a long right-hand tail.

Since the VaR is a loss, and since the loss is the difference between P_t (which is random) and P_{t-1} (which we can take here as given), then the VaR itself has the same distribution as P_t. Normally distributed geometric returns imply that the VaR is lognormally distributed.

If we proceed as we did earlier with the arithmetic return, we begin by deriving the critical value of R, R^*, such that the probability that $R > R^*$ is α:

$$R^* = \mu_R - \sigma_R z_\alpha \tag{1.15}$$

BOX 1.1 THE LOGNORMAL DISTRIBUTION

A random variate X is said to be lognormally distributed if the natural log of X is normally distributed. The lognormal distribution can be specified in terms of the mean and standard deviation of $\ln X$. Call these parameters μ and σ. The lognormal is often also represented in terms of m and σ, where m is the median of x, and $m = \exp(\mu)$.

The pdf of X can be written:

$$\phi(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\log(x) - \mu}{\sigma}\right)^2\right\}$$

for $x > 0$. Thus, the lognormal pdf is only defined for positive values of x and is skewed to the right as in Figure 1.6.

Let $\omega = \exp(\sigma^2)$ for convenience. The mean and variance of the lognormal can be written as:

$$\text{mean} = m\exp(\sigma^2/2) \quad \text{and} \quad \text{variance} = m^2\omega(\omega - 1)$$

Turning to higher moments, the skewness of the lognormal is

$$skewness = (\omega + 2)(\omega - 1)^{1/2}$$

and is always positive, which confirms the lognormal has a long right-hand tail. The kurtosis of the lognormal is

$$kurtosis = \omega^4 + 2\omega^3 + 3\omega^2 - 3$$

and therefore varies from a minimum of (just over) 3 to a potentially large value depending on the value of s.

We then use the definition of the geometric return to unravel the critical value of P^* (i.e., the value of P_t corresponding to a loss equal to our VaR), and thence infer our VaR:

$$R^* = \ln(P^*/P_{t-1}) = \ln P^* - \ln P_{t-1}$$
$$\Rightarrow \ln P^* = R^* + \ln P_{t-1}$$
$$\Rightarrow P^* = P_{t-1}\exp[R^*] = P_{t-1}\exp[\mu_R - \sigma_R z_\alpha]$$
$$\Rightarrow \alpha VaR = P_{t-1} - P^* = P_{t-1}(1 - \exp[\mu_R - \sigma_R z_\alpha]) \tag{1.16}$$

This gives us the lognormal VaR, which is consistent with normally distributed geometric returns.

The lognormal VaR is illustrated in Figure 1.7, based on the standardised (but typically unrealistic) assumptions that $\mu_R = 0$, $\sigma_R = 1$, and $P_{t-1} = 1$. In this case, the VaR at the 95% confidence level is 0.807. The figure also shows that the distribution of L/P is a reflection of the distribution of P_t shown earlier in Figure 1.6.

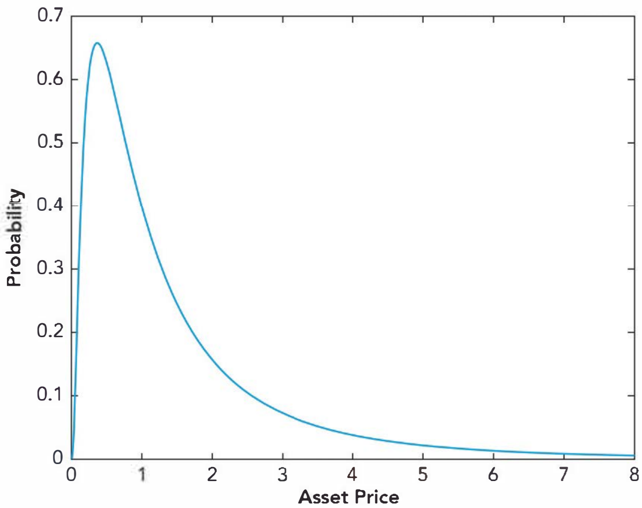

Figure 1.6 **A lognormally distributed asset price.**

Note: Estimated using the 'lognpdf' function in the Statistics Toolbox.

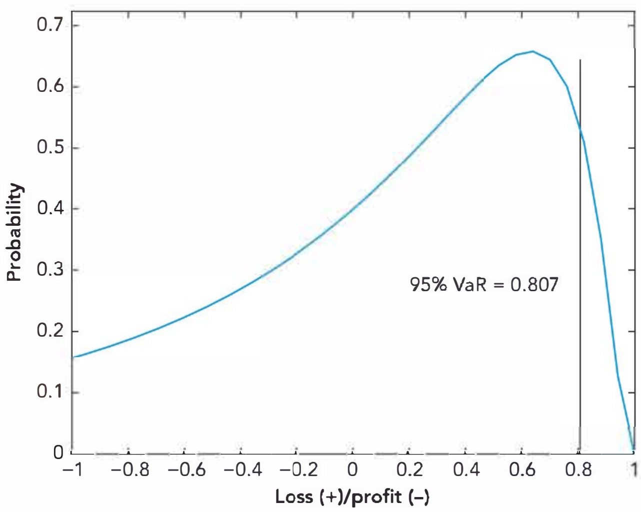

Figure 1.7 **Lognormal VaR.**

Note: Estimated assuming the mean and standard deviation of geometric returns are 0 and 1, and for an initial investment of 1. The figure is produced using the 'lognormalvarfigure' function.

Example 1.4 Lognormal VaR

Suppose that geometric returns R_t over some period are distributed as normal with mean 0.05, standard deviation 0.20, and we have a portfolio currently worth 1. Then (by Equation (1.16)) the 95% VaR is $1 - \exp(0.05 - 0.20 \times 1.645) = 0.244$.

The corresponding 99% VaR is $1 - \exp(0.05 - 0.20 \times 2.326) = 0.340$. Observe that these VaRs are quite close to those obtained in Example 1.3, where the arithmetic return parameters were the same as the geometric return parameters assumed here.

Example 1.5 Lognormal VaR vs Normal VaR

Suppose that we make the empirically not too unrealistic assumptions that the mean and volatility of annualised returns are 0.10 and 0.40. We are interested in the 95% VaR at the 1-day holding period for a portfolio worth USD 1. Assuming 250 trading days to a year, the daily return has a mean $0.1/250 = 0.00040$ and standard deviation $0.40/\sqrt{250} = 0.0253$. The normal 95% VaR is $-0.0004 + 0.0253 \times 1.645 = 0.0412$. If we assume a lognormal, then the 95% VaR is $1 - \exp(0.0004 - 0.0253 \times 1.645) = 0.0404$. The normal VaR is 4.12% and the lognormal VaR is 4.04% of the value of the portfolio. This illustrates that normal and lognormal VaRs are much the same if we are dealing with short holding periods and realistic return parameters.

1.4 ESTIMATING COHERENT RISK MEASURES

Estimating Expected Shortfall

We turn now to the estimation of coherent risk measures, and the easiest of these to estimate is the expected shortfall (ES). The ES is the probability-weighted average of tail losses, and a normal ES is illustrated in Figure 1.8. In this case, the 95% ES is 2.063, corresponding to our earlier normal 95% VaR of 1.645.

The fact that the ES is a probability-weighted average of tail losses suggests that we can estimate ES as an average of 'tail VaRs'.[4] The easiest way to implement this approach is to slice the tail into a large number *n* of slices, each of which has the same probability mass, estimate the VaR associated with each slice, and take the ES as the average of these VaRs.

To illustrate the method, suppose we wish to estimate a 95% ES on the assumption that losses are normally distributed with mean 0 and standard deviation 1. In practice, we would use a

[4] The obvious alternative is to seek a 'closed-form' solution, which we could use to estimate the ES, but ES formulas seem to be known only for a limited number of parametric distributions (e.g., elliptical, including normal, and generalised Pareto distributions), whereas the 'average-tail-VaR' method is easy to implement and can be applied to any 'well-behaved' ESs that we might encounter, parametric or otherwise.

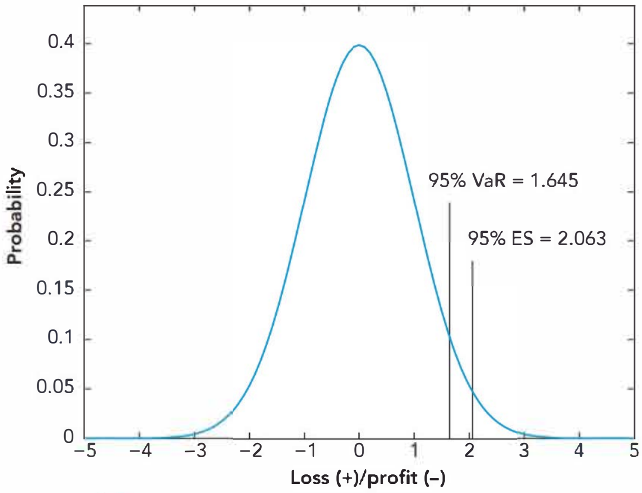

Figure 1.8 Normal VaR and ES.

Note: Estimated with the mean and standard deviation of P/L equal to 0 and 1 respectively, using the 'normalesfigure' function.

high value of n and carry out the calculations on a spreadsheet or using appropriate software. However, to show the procedure manually, let us work with a very small n value of 10. This value gives us 9 (i.e., $n - 1$) tail VaRs, or VaRs at confidence levels in excess of 95%. These VaRs are shown in Table 1.1, and vary from 1.6954 (for the 95.5% VaR) to 2.5758 (for the 99.5% VaR). Our estimated ES is the average of these VaRs, which is 2.0250.

Table 1.1 Estimating ES as a Weighted Average of Tail VaRs

Confidence Level	Tail VaR
95.5%	1.6954
96.0%	1.7507
96.5%	1.8119
97.0%	1.8808
97.5%	1.9600
98.0%	2.0537
98.5%	2.1701
99.0%	2.3263
99.5%	2.5738
Average of tail VaRs	2.0250

Note: VaRs estimated assuming the mean and standard deviation of losses are 0 and 1, using the 'normalvar' function in the MMR Toolbox.

Of course, in using this method for practical purposes, we would want a value of n large enough to give accurate results. To give some idea of what this might be, Table 1.2 reports some alternative ES estimates obtained using this procedure with varying values of n. These results show that the estimated ES rises with n, and gradually converges to the true value of 2.063. These results also show that our ES estimation procedure seems to be reasonably accurate even for quite small values of n. Any decent computer should therefore be able to produce accurate ES estimates quickly in real time.

Estimating Coherent Risk Measures

Other coherent risk measures can be estimated using modifications of this 'average VaR' method. Recall that a coherent risk measure is a weighted average of the quantiles (denoted by q_p) of our loss distribution:

$$M_\phi = \int_0^1 \phi(p) q_p dp \tag{1.17}$$

where the weighting function or risk-aversion function $\phi(p)$ is specified by the user. The ES gives all tail-loss quantiles an equal weight, and other quantiles a weight of 0. Thus the ES is a special case of M_ϕ obtained by setting $\phi(p)$ to the following:

$$\phi(p) = \begin{cases} 0 & \text{if } p < \alpha \\ 1/(1-\alpha) & \text{if } p \geq \alpha \end{cases} \tag{1.18}$$

Table 1.2 ES Estimates as a Function of the Number of Tail Slices

Number of Tail Slices (n)	ES
10	2.0250
25	2.0433
50	2.0513
100	2.0562
250	2.0597
500	2.0610
1000	2.0618
2500	2.0623
5000	2.0625
10 000	2.0626
True value	2.0630

Note: VaRs estimated assuming the mean and standard deviation of losses are 0 and 1.

The more general coherent risk measure, M_ϕ, involves a potentially more sophisticated weighting function $\phi(p)$. We can therefore estimate any of these measures by replacing the equal weights in the 'average VaR' algorithm with the $\phi(p)$ weights appropriate to the risk measure being estimated.

To show how this might be done, suppose we have the exponential weighting function:

$$\phi_\gamma(p) = \frac{e^{-(1-p)/\gamma}}{\gamma(1 - e^{-1/\gamma})} \quad (1.19)$$

and we believe that we can represent the degree of our risk-aversion by setting $\gamma = 0.05$. To illustrate the procedure manually, we continue to assume that losses are standard normally distributed and we set $n = 10$ (i.e., we divide the complete losses density function into 10 equal-probability slices). With $n = 10$, we have $n - 1 = 9$ (i.e., $n - 1$) loss quantiles or VaRs spanning confidence levels from 0.1 to 0.90. These VaRs are shown in the second column of Table 1.3, and vary from −1.2816 (for the 10% VaR) to 1.2816 (for the 90% VaR). The third column shows the $\phi(p)$ weights corresponding to each confidence level, and the fourth column shows the products of each VaR and corresponding weight. Our estimated exponential spectral risk measure is the $\phi(p)$-weighted average of the VaRs, and is therefore equal to 0.4228.

As when estimating the ES earlier, when using this type of routine in practice we would want a value of n large enough to give accurate results. Table 1.4 reports some alternative estimates obtained using this procedure with increasing values of n. These results show that the estimated risk measure rises with n, and gradually converges to a value in the region of about 1.854. The estimates in this table indicate that we may need a considerably larger value of n than we did earlier to get results of the same level of accuracy. Even so, a good computer should still be able to produce accurate estimates of spectral risk measures fairly quickly.

When estimating ES or more general coherent risk measures in practice, it also helps to have some guidance on how to choose the value of n. Granted that the estimate does eventually converge to the true value as n gets large, one useful approach is to start with some small value of n, and then double n repeatedly until we feel the estimates have settled down sufficiently. Each time we do so, we halve the width of the discrete slices, and we can monitor how this 'halving' process affects our estimates. This suggests that we look at the 'halving error' ε_n given by:

$$\varepsilon_n = \hat{M}^{(n)} - \hat{M}^{(n/2)} \quad (1.20)$$

where $\hat{M}^{(n)}$ is our estimated risk measure based on n slices. We stop doubling n when ε_n falls below some tolerance level that indicates an acceptable level of accuracy. The process is

Table 1.3 Estimating Exponential Spectral Risk Measure as a Weighted Average of VaRs

Confidence Level (α)	α VaR	Weight $\phi(\alpha)$	$\phi(\alpha) \times \alpha$ VaR
10%	−1.2816	0	0.0000
20%	−0.8416	0	0.0000
30%	−0.5244	0	0.0000
40%	−0.2533	0.0001	0.0000
50%	0	0.0009	0.0000
60%	0.2533	0.0067	0.0017
70%	0.5244	0.0496	0.0260
80%	0.8416	0.3663	0.3083
90%	1.2816	2.7067	3.4689
Risk measure = mean ($\phi(\alpha)$ times αVaR) =			0.4226

Note: VaRs estimated assuming the mean and standard deviation of losses are 0 and 1, using the 'normalvar' function in the MMR Toolbox. The weights $\phi(\alpha)$ are given by the exponential function (Equation (1.19)) with $\gamma = 0.05$.

Table 1.4 Estimates of Exponential Spectral Coherent Risk Measure as a Function of the Number of Tail Slices

Number of Tail Slices	Estimate of Exponential Spectral Risk Measure
10	0.4227
50	1.3739
100	1.5853
250	1.7338
500	1.7896
1000	1.8197
2500	1.8392
5000	1.8461
10 000	1.8498
50 000	1.8529
100 000	1.8533
500 000	1.8536

Note: VaRs estimated assuming the mean and standard deviation of losses are 0 and 1, using the 'normalvar' function in the MMR Toolbox. The weights $\phi(\alpha)$ are given by the exponential function (Equation (1.19)) with $\gamma = 0.05$.

Table 1.5 Estimated Risk Measures and Halving Errors

Number of Tail Slices	Estimated Spectral Risk Measure	Halving Error
100	1.5853	0.2114
200	1.7074	0.1221
400	1.7751	0.0678
800	1.8120	0.0368
1600	1.8317	0.0197
3200	1.8422	0.0105
6400	1.8477	0.0055
12 800	1.8506	0.0029
25 600	1.8521	0.0015
51 200	1.8529	0.0008

Note: VaRs estimated assuming the mean and standard deviation of losses are 0 and 1, using the 'normalvar' function in the MMR Toolbox. The weights $\phi(\alpha)$ are given by the exponential function (Equation (1.19)) with $\gamma = 0.05$.

shown in Table 1.5. Starting from an arbitrary value of 100, we repeatedly double n (so it becomes 200, 400, 800, etc.). As we do so, the estimated risk measure gradually converges, and the halving error gradually falls. So, for example, for $n = 6400$, the estimated risk measure is 1.8477, and the halving error is 0.0055. If we double n to 12,800, the estimated risk measure becomes 1.8506, and the halving error falls to 0.0029, and so on.

However, this 'weighted average quantile' procedure is rather crude, and (bearing in mind that the risk measure (Equation (1.17)) involves an integral) we can in principle expect to get substantial improvements in accuracy if we resorted to more 'respectable' numerical integration or quadrature methods. This said, the crude 'weighted average quantile' method actually seems to perform well for spectral exponential risk measures when compared against some of these alternatives, so one is not necessarily better off with the more sophisticated methods.[5]

[5] There is an interesting reason for this: the spectral weights give the highest loss the highest weight, whereas the quadrature methods such as the trapezoidal and Simpson's rules involve algorithms in which the two most extreme quantiles have their weights specifically cut, and this undermines the accuracy of the algorithm relative to the crude approach. However, there are ways round these sorts of problems, and in principle versions of the sophisticated approaches should give better results.

Thus, the key to estimating any coherent risk measure is to be able to estimate quantiles or VaRs: the coherent risk measures can then be obtained as appropriately weighted averages of quantiles. From a practical point of view, this is extremely helpful as all the building blocks that go into quantile or VaR estimation—databases, calculation routines, etc.—are exactly what we need for the estimation of coherent risk measures as well. If an institution already has a VaR engine, then that engine needs only small adjustments to produce estimates of coherent risk measures: indeed, in many cases, all that needs changing is the last few lines of code in a long data processing system. The costs of switching from VaR to more sophisticated risk measures are therefore very low.

1.5 ESTIMATING THE STANDARD ERRORS OF RISK MEASURE ESTIMATORS

We should always bear in mind that any risk measure estimates that we produce are just that—estimates. We never know the true value of any risk measure, and an estimate is only as good as its precision: if a risk measure is very imprecisely estimated, then the estimator is virtually worthless, because its imprecision tells us that true value could be almost anything; on the other hand, if we know that an estimator is fairly precise, we can be confident that the true value is fairly close to the estimate, and the estimator has some value. Hence, we should always seek to supplement any risk estimates we produce with some indicator of their precision. This is a fundamental principle of good risk measurement practice.

We can evaluate the precision of estimators of risk measures by means of their standard errors, or (generally better) by producing confidence intervals for them. In this chapter we focus on the more basic indicator, the standard error of a risk measure estimator.

Standard Errors of Quantile Estimators

We first consider the standard errors of quantile (or VaR) estimators. Following Kendall and Stuart,[6] suppose we have a distribution (or cumulative density) function $F(x)$, which might be a parametric distribution function or an empirical

[6] Kendall and Stuart (1972), pp. 251–252.

distribution function (i.e., a cumulative histogram) estimated from real data. Its corresponding density or relative-frequency function is $f(x)$. Suppose also that we have a sample of size n, and we select a bin width h. Let dF be the probability that $(k - 1)$ observations fall below some value $q - h/2$, that one observation falls in the range $q \pm h/2$, and that $(n - k)$ observations are greater than $q + h/2$. dF is proportional to

$$\{F(q)\}^{k-1} f(q)dq\{1 - F(q)\}^{n-k} \tag{1.21}$$

This gives us the frequency function for the quantile q not exceeded by a proportion k/n of our sample, i.e., the $100(k/n)$th percentile.

If this proportion is p, Kendall and Stuart show that Equation (1.21) is approximately equal to $p^{np}(1 - p)^{n(1-p)}$ for large values of n. If ε is a very small increment to p, then

$$p^{np}(1 - p)^{n(1-p)} \approx (p + \varepsilon)^{np}(1 - p - \varepsilon)^{n(1-p)} \tag{1.22}$$

Taking logs and expanding, Equation (1.22) is itself approximately

$$(p + \varepsilon)^{np}(1 - p - \varepsilon)^{n(1-p)} \approx \frac{n\varepsilon^2}{2p(1 - p)} \tag{1.23}$$

which implies that the distribution function dF is approximately proportional to

$$\exp\left(\frac{-n\varepsilon^2}{2p(1 - p)}\right) \tag{1.24}$$

Integrating this out,

$$dF = \frac{1}{\sqrt{2\pi}\sqrt{p(1 - p)/n}} \exp\left(\frac{-n\varepsilon^2}{2p(1 - p)}\right) d\varepsilon \tag{1.25}$$

which tells us that ε is normally distributed in the limit with variance $p(1 - p)/n$. However, we know that $d\epsilon = dF(q) = f(q)dq$, so the variance of q is

$$\mathrm{var}(q) \approx \frac{p(1 - p)}{n[f(q)]^2} \tag{1.26}$$

This gives us an approximate expression for the variance, and hence its square root, the standard error, of a quantile estimator q.

This expression shows that the quantile standard error depends on p, the sample size n and the pdf value $f(q)$. The way in which the (normal) quantile standard errors depend on these parameters is apparent from Figure 1.9. This shows that:

- The standard error falls as the sample size n rises.
- The standard error rises as the probabilities become more extreme and we move further into the tail—hence, the more extreme the quantile, the less precise its estimator.

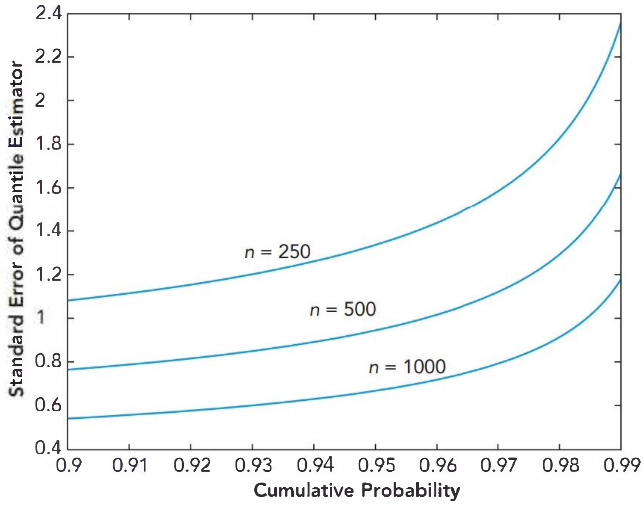

Figure 1.9 **Standard errors of quantile estimators.**

Note: Based on random samples of size n drawn from a standard normal distribution. The bin width h is set to 0.1.

In addition, the quantile standard error depends on the probability density function $f(.)$—so the choice of density function can make a difference to our estimates—and also on the bin width h, which is essentially arbitrary.

The standard error can be used to construct confidence intervals around our quantile estimates in the usual textbook way. For example, a 90% confidence interval for a quantile q is given by

$$[q - 1.645\ \mathrm{se}(q),\ q + 1.645\ \mathrm{se}(q)]$$

$$= \left[q - 1.645\frac{\sqrt{p(1 - p)/n}}{f(q)},\ q + 1.645\frac{\sqrt{p(1 - p)/n}}{f(q)}\right] \tag{1.27}$$

Example 1.6 Obtaining VaR Confidence Intervals Using Quantile Standard Errors

Suppose we wish to estimate the 90% confidence interval for a 95% VaR estimated on a sample of size of $n = 1000$ to be drawn from a standard normal distribution, based on an assumed bin width $h = 0.1$.

We know that the 95% VaR of a standard normal is 1.645. We take this to be q in Equation (1.27), and we know that q falls in the bin spanning $1.645 \pm 0.1/2 = [1.595, 1.695]$. The probability of a loss exceeding 1.695 is 0.045, and this is also equal to p, and the probability of profit or a loss less than 1.595 is 0.9446. Hence $f(q)$, the probability mass in the q range, is $1 - 0.0450 - 0.9446 = 0.0104$. We now plug the relevant values into Equation (1.27) to obtain the 90% confidence interval for the VaR:

$$\left[1.645 - 1.645\frac{\sqrt{0.045(1-0.045)/1000}}{0.0104},\ 1.645 + 1.645\frac{\sqrt{0.045(1-0.045)/1000}}{0.0104}\right] = [0.6081, 2.6819]$$

This is a wide confidence interval, especially when compared to the OS and bootstrap confidence intervals.

The confidence interval narrows if we take a wider bin width, so suppose that we now repeat the exercise using a bin width $h = 0.2$, which is probably as wide as we can reasonably go with these data. q now falls into the range $1.645 \pm 0.2/2 = [1.545, 1.745]$. p, the probability of a loss exceeding 1.745, is 0.0405, and the probability of profit or a loss less than 1.545 is 0.9388. Hence $f(q) = 1 - 0.0405 - 0.9388 = 0.0207$. Plugging these values into Equation (1.27) now gives us a new estimate of the 90% confidence interval:

$$\left[1.645 - 1.645\frac{\sqrt{0.0405(1-0.0405)/1000}}{0.0207},\ 1.645 + 1.645\frac{\sqrt{0.0405(1-0.0405)/1000}}{0.0207}\right] = [1.1496, 2.1404]$$

This is still a rather wide confidence interval.

This example illustrates that although we can use quantile standard errors to estimate VaR confidence intervals, the intervals can be wide and also sensitive to the arbitrary choice of bin width.

The quantile-standard-error approach is easy to implement and has some plausibility with large sample sizes. However, it also has weaknesses relative to other methods of assessing the precision of quantile (or VaR) estimators—it relies on asymptotic theory and requires large sample sizes; it can produce imprecise estimators, or wide confidence intervals; it depends on the arbitrary choice of bin width; and the symmetric confidence intervals it produces are misleading for extreme quantiles whose 'true' confidence intervals are asymmetric reflecting the increasing sparsity of extreme observations as we move further out into the tail.

Standard Errors in Estimators of Coherent Risk Measures

We now consider standard errors in estimators of coherent risk measures. One of the first studies to examine this issue (Yamai and Yoshiba (2001b) did so by investigating the relative accuracy of VaR and ES estimators for Lévy distributions with varying α stability parameters. Their results suggested that VaR and ES estimators had comparable standard errors for near-normal Lévy distributions, but the ES estimators had much bigger standard errors for particularly heavy-tailed distributions. They explained this finding by saying that as tails became heavier, ES estimators became more prone to the effects of large but infrequent losses. This finding suggests the depressing conclusion that the presence of heavy tails might make ES estimators in general less accurate than VaR estimators.

Fortunately, there are grounds to think that such a conclusion might be overly pessimistic. For example, Inui and Kijima (2003) present alternative results showing that the application of a Richardson extrapolation method can produce ES estimators that are both unbiased and have comparable standard errors to VaR estimators.[7] Acerbi (2004) also looked at this issue and, although he confirmed that tail heaviness did increase the standard errors of ES estimators relative to VaR estimators, he concluded that the relative accuracies of VaR and ES estimators were roughly comparable in empirically realistic ranges.

However, the standard error of any estimator of a coherent risk measure will vary from one situation to another, and the best practical advice is to get into the habit of always estimating the standard error whenever one estimates the risk measure itself. Estimating the standard error of an estimator of a coherent risk measure is also relatively straightforward. One way to do so starts from recognition that a coherent risk measure is an L-estimator (i.e., a weighted average of order statistics), and L-estimators are asymptotically normal. If we take N discrete points in the density function, then as N gets large the variance of the estimator of the coherent risk measure (Equation (1.17)) is approximately

$$\sigma(M_\phi^{(N)}) \rightarrow \frac{2}{N}\int_{p<q} \phi(p)\phi(q)\frac{p(1-q)}{f(F^{-1}(p))f(F^{-1}(q))}dpdq = \frac{2}{N}\int_{x<y} \phi(F(x))\phi(F(y))F(x)(1-F(y))dxdy \tag{1.28}$$

and this can be computed numerically using a suitable numerical integration procedure. Where the risk measure is the ES, the standard error becomes

$$\sigma(ES^{(N)}) \rightarrow \frac{1}{N\alpha^2}\int_0^{F^{-1}(\alpha)}\int_0^{F^{-1}(\alpha)} [\min(F(x), F(y)) - F(x)F(y)]dxdy \tag{1.29}$$

and used in conjunction with a suitable numerical integration method, this gives good estimates even for relatively low values

[7] See Inui and Kijima (2003).

of N.[8] If we wish to obtain confidence intervals for our risk measure estimators, we can make use of the asymptotic normality of these estimators to apply textbook formulas (e.g., such as Equation (1.27)) based on the estimated standard errors and centred around a 'good' best estimate of the risk measure.

An alternative approach to the estimation of standard errors for estimators of coherent risk measures is to apply a bootstrap: we bootstrap a large number of estimators from the given distribution function (which might be parametric or non-parametric, e.g., historical); and we estimate the standard error of the sample of bootstrapped estimators. Even better, we can also use a bootstrapped sample of estimators to estimate a confidence interval for our risk measure.

1.6 THE CORE ISSUES: AN OVERVIEW

Before proceeding to more detailed issues, it might be helpful to pause for a moment to take an overview of the structure, as it were, of the subject matter itself. This is very useful, as it gives the reader a mental frame of reference within which the 'detailed' material that follows can be placed. Essentially, there are three core issues, and all the material that follows can be related to these. They also have a natural sequence, so we can think of them as providing a roadmap that leads us to where we want to be.

Which risk measure? The first and most important is to choose the type of risk measure: do we want to estimate VaR, ES, etc.? This is logically the first issue, because we need to know *what* we are trying to estimate before we start thinking about *how* we are going to estimate it.

Which level? The second issue is the *level* of analysis. Do we wish to estimate our risk measure at the level of the portfolio as a whole or at the level of the individual positions in it? The former would involve us taking the portfolio as our basic unit of analysis (i.e., we take the portfolio to have a specified composition, which is taken as given for the purposes of our analysis), and this will lead to a *univariate* stochastic analysis. The alternative is to work from the position level, and this has the advantage of allowing us to accommodate changes in the portfolio composition within the analysis itself. The disadvantage is that we then need a *multivariate* stochastic framework, and this is considerably more difficult to handle: we have to get to grips with the problems posed by variance–covariance matrices, copulas, and so on, all of which are avoided if we work at the portfolio level. There is thus a trade-off: working at the portfolio level is more limiting, but easier, while working at the position level gives us much more flexibility, but can involve much more work.

Which method? Having chosen our risk measure and level of analysis, we then choose a suitable estimation method. To decide on this, we would usually think in terms of the classic 'VaR trinity':

- Non-parametric methods
- Parametric methods
- Monte Carlo simulation methods

Each of these involves some complex issues.

[8] See Acerbi (2004, pp. 200–201).

1.7 APPENDIX

Preliminary Data Analysis

When confronted with a new data set, we should *never* proceed straight to estimation without some preliminary analysis to get to know our data. Preliminary data analysis is useful because it gives us a feel for our data, and because it can highlight problems with our data set. Remember that we never really know where our data come from, so we should always be a little wary of any new data set, regardless of how reputable the source might appear to be. For example, how do you know that a clerk hasn't made a mistake somewhere along the line in copying the data and, say, put a decimal point in the wrong place? The answer is that you don't, and never can. Even the most reputable data providers provide data with errors in them, however careful they are. Everyone who has ever done any empirical work will have encountered such problems at some time or other: the bottom line is that real data must always be viewed with a certain amount of suspicion.

Such preliminary analysis should consist of at least the first two and preferably all three of the following steps:

- The first and by far the most important step is to eyeball the data to see if they 'look right'—or, more to the point, we should eyeball the data to see if anything looks *wrong*. Does the pattern of observations look right? Do any observations stand out as questionable? And so on. The interocular trauma test is the most important test ever invented and also the easiest to carry out, and we should always perform it on any new data set.
- We should plot our data on a histogram and estimate the relevant summary statistics (i.e., mean, standard deviation, skewness, kurtosis, etc.). In risk measurement, we are particularly interested in any non-normal features of our data: skewness, excess kurtosis, out-liers in our data, and the like.

We should therefore be on the lookout for any evidence of non-normality, and we should take any such evidence into account when considering whether to fit any parametric distribution to the data.

- Having done this initial analysis, we should consider what kind of distribution might fit our data, and there are a number of useful diagnostic tools available for this purpose, the most popular of which are QQ plots—plots of empirical quantiles against their theoretical equivalents.

Plotting the Data and Evaluating Summary Statistics

To get to know our data, we should first obtain their histogram and see what might stand out. Do the data look normal, or non-normal? Do they show one pronounced peak, or more than one? Do they seem to be skewed? Do they have fat tails or thin tails? Are there outliers? And so on.

As an example, Figure 1.10 shows a histogram of 100 random observations. In practice, we would usually wish to work with considerably longer data sets, but a data set this small helps to highlight the uncertainties one often encounters in practice. These observations show a dominant peak in the centre, which suggests that they are probably drawn from a unimodal distribution. On the other hand, there may be a negative skew, and there are some large outlying observations on the extreme left of the distribution, which might indicate fat tails on at least the left-hand side. In fact, these particular observations are drawn from a Student-*t* distribution with 5 degrees of freedom, so in this case we know that the underlying true distribution is unimodal, symmetric and heavy tailed. However, we would not know this in a situation with 'real' data, and it is precisely because we do not know the distributions of real-world data sets that preliminary analysis is so important.

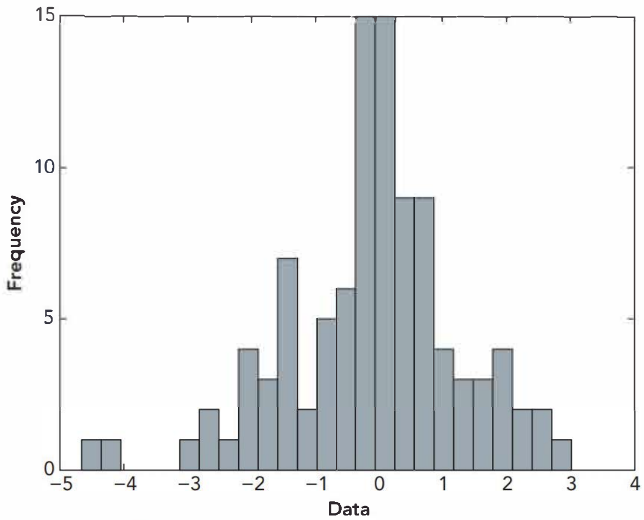

Figure 1.10 **A histogram.**

Note: Data are 100 observations randomly drawn from a Student-*t* with 5 degrees of freedom.

Some summary statistics for this data set are shown in Table 1.6. The sample mean (−0.099) and the sample mode differ somewhat (−0.030), but this difference is small relative to the sample standard deviation (1.363). However, the sample skew (−0.503) is somewhat negative and the sample kurtosis (3.985) is a little bigger than normal. The sample minimum (−4.660) and the sample maximum (3.010) are also not symmetric about the sample mean or mode, which is further evidence of asymmetry. If we encountered these results with 'real' data, we would be concerned about possible skewness and kurtosis. However, in this hypothetical case we know that the sample skewness is merely a product of sample variation, because we happen to know that the data are drawn from a symmetric distribution.

Depending on the context, we might also seriously consider carrying out some formal tests. For example, we might test whether the sample parameters (mean, standard deviation, etc.) are consistent with what we might expect under a null hypothesis (e.g., such as normality).

The underlying principle is very simple: since we *never* know the true distribution in practice, all we ever have to work with are *estimates* based on the *sample* at hand; it therefore behoves us to make the best use of the data we have, and to extract as much information as possible from them.

QQ Plots

Having done our initial analysis, it is often good practice to ask what distribution might fit our data, and a very useful device for identifying the distribution of our data is a quantile–quantile or QQ plot—a plot of the quantiles of the empirical distribution against those of some specified distribution. The shape of the QQ plot tells us a lot about how the empirical distribution compares to the specified one. In particular, if the QQ plot is linear, then the specified distribution fits the data, and we have identified the distribution to which our data belong. In addition, departures of the QQ from linearity in the tails can tell us whether the tails of our empirical distribution are fatter, or thinner, than the tails of the reference distribution to which it is being compared.

Table 1.6 Summary Statistics

Parameter	Value
Mean	−0.099
Mode	−0.030
Standard deviation	1.363
Skewness	−0.503
Kurtosis	3.985
Minimum	−4.660
Maximum	3.010
Number of observations	100

Note: Data are the same observations shown in Figure 1.10.

To illustrate, Figure 1.11 shows a QQ plot for a data sample drawn from a normal distribution, compared to a reference distribution that is also normal. The QQ plot is obviously close to linear: the central mass observations fit a linear QQ plot very closely, and the extreme tail observations somewhat less so. However, there is no denying that the overall plot is approximately linear. Figure 1.11 is a classic example of a QQ plot in which the empirical distribution matches the reference population.

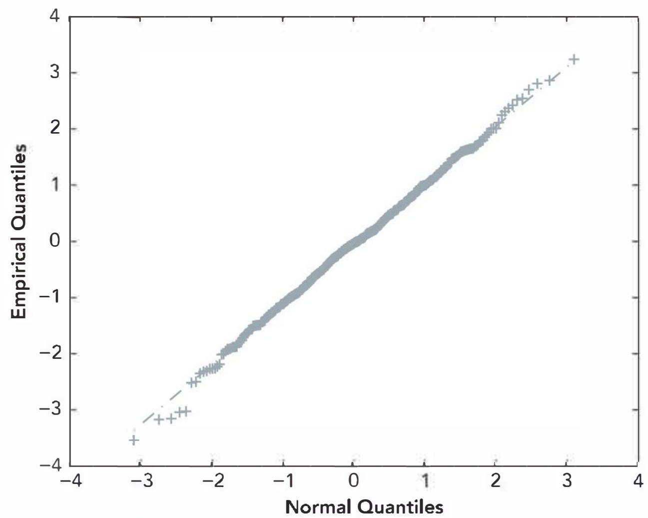

Figure 1.11 **QQ plot: normal sample against normal reference distribution.**

Note: The empirical sample is a random sample of 500 observations drawn from a standard normal. The reference distribution is standard normal.

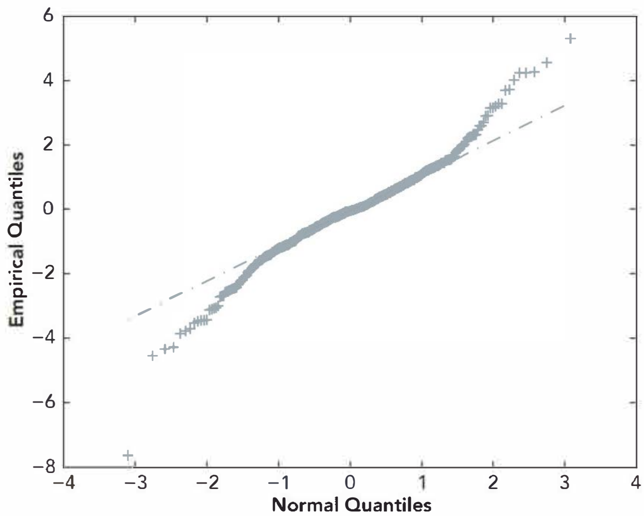

Figure 1.12 **QQ plot: *t* sample against normal reference distribution.**

Note: The empirical sample is a random sample of 500 observations drawn from Student-*t* with 5 degrees of freedom. The reference distribution is standard normal.

By contrast, Figure 1.12 shows a good example of a QQ plot where the empirical distribution does not match the reference population. In this case, the data are drawn from a Student-*t* with 5 degrees of freedom, but the reference distribution is standard normal. The QQ plot is now clearly non-linear: although the central mass observations are close to linear, the tails show steeper slopes indicative of the tails being heavier than those of the reference distribution.

A QQ plot is useful in a number of ways. First, as noted already, if the data are drawn from the reference population, then the QQ plot should be linear. This remains true if the data are drawn from some linear transformation of the reference distribution (i.e., are drawn from the same distribution but with different location and scale parameters). We can therefore use a QQ plot to form a tentative view of the distribution from which our data might be drawn: we specify a variety of alternative distributions, and construct QQ plots for each. Any reference distributions that produce non-linear QQ plots can then be dismissed, and any distribution that produces a linear QQ plot is a good candidate distribution for our data.

Second, because a linear transformation in one of the distributions in a QQ plot merely changes the intercept and slope of the QQ plot, we can use the intercept and slope of a linear QQ plot to give us a rough idea of the location and scale parameters of our sample data. For example, the reference distribution in

Figure 1.11 is a standard normal. The linearity of the QQ plot in this figure suggests that the data are normal, as mentioned already, but Figure 1.11 also shows that the intercept and slope are approximately 0 and 1 respectively, and this indicates that the data are drawn from a standard normal, and not just any normal. Such rough approximations give us a helpful yardstick against which we can judge more 'sophisticated' estimates of location and scale, and also provide useful initial values for iterative algorithms.

Third, if the empirical distribution has heavier tails than the reference distribution, the QQ plot will have steeper slopes at its tails, even if the central mass of the empirical observations are approximately linear. Figure 1.12 is a good case in point. A QQ plot where the tails have slopes different than the central mass is therefore suggestive of the empirical distribution having heavier, or thinner, tails than the reference distribution.

Finally, a QQ plot is good for identifying outliers (e.g., observations contaminated by large errors): such observations will stand out in a QQ plot, even if the other observations are broadly consistent with the reference distribution.[9]

[9] Another useful tool, especially when dealing with the tails, is the mean excess function (MEF): the expected amount by which a random variable X exceeds some threshold u, given that $X > u$. The usefulness of the MEF stems from the fact that each distribution has its own distinctive MEF. A comparison of the empirical MEF with the theoretical MEF associated with some specified distribution function therefore gives us an indication of whether the chosen distribution fits the tails of our empirical distribution. However, the results of MEF plots need to be interpreted with some care, because data observations become more scarce as X gets larger. For more on these and how they can be used, see Embrechts et. al. (1997, Chapters 3.4 and 6.2).

2

Non-Parametric Approaches

Learning Objectives

After completing this reading, you should be able to:

- Apply the bootstrap historical simulation approach to estimate coherent risk measures.
- Describe historical simulation using non-parametric density estimation.
- Compare and contrast the age-weighted, the volatility-weighted, the correlation-weighted, and the filtered historical simulation approaches.
- Identify advantages and disadvantages of non-parametric estimation methods.

Excerpt is Chapter 4 of Measuring Market Risk, Second Edition, by Kevin Dowd.