19

Risk Management

19.1 The Need for Risk Management

The financial world has always been risky, and financial innovations such as the development of derivatives markets and the packaging of mortgages have now made risk management more important than ever but also more difficult.

There are many different types of risk. *Market risk* is due to changes in prices. *Credit risk* is the danger that a counterparty does not meet contractual obligations, for example, that interest or principal on a bond is not paid. *Liquidity risk* is the potential extra cost of liquidating a position because buyers are difficult to locate. *Operational risk* is due to fraud, mismanagement, human errors, and similar problems.

Early attempts to measure risk such as duration analysis, discussed in Section 3.8.1 and used to estimate the market risk of fixed income securities, were somewhat primitive and of only limited applicability. In contrast, value-at-risk (VaR) and expected shortfall (ES) are widely used because they can be applied to all types of risks and securities, including complex portfolios.

VaR uses two parameters, the time horizon and the confidence level, which are denoted by T and $1 - \alpha$, respectively. Given these, the VaR is a bound such that the loss over the horizon is less than this bound with probability equal to the confidence coefficient. For example, if the horizon is one week, the confidence coefficient is 99% (so $\alpha = 0.01$), and the VaR is \$5 million, then there is only a 1% chance of a loss exceeding \$5 million over the next week. We sometimes use the notation $\mathrm{VaR}(\alpha)$ or $\mathrm{Var}(\alpha, T)$ to indicate the dependence of VaR on α or on both α and the horizon T. Usually, $\mathrm{VaR}(\alpha)$ is used with T being understood.

If $\mathcal{L}$ is the loss over the holding period T, then $\mathrm{VaR}(\alpha)$ is the αth upper quantile of $\mathcal{L}$. Equivalently, if $\mathcal{R} = -\mathcal{L}$ is the revenue, then $\mathrm{VaR}(\alpha)$ is minus the αth quantile of $\mathcal{R}$. For continuous loss distributions, $\mathrm{VaR}(\alpha)$ solves

$$P\{\mathcal{L} > \mathrm{VaR}(\alpha)\} = P\{\mathcal{L} \geq \mathrm{VaR}(\alpha)\} = \alpha, \tag{19.1}$$

and for any loss distribution, continuous or not,

D. Ruppert, *Statistics and Data Analysis for Financial Engineering*, Springer Texts in Statistics,
DOI 10.1007/978-1-4419-7787-8_19,

$$\text{VaR}(\alpha) = \inf\{x : P(\mathcal{L} > x) \leq \alpha\}. \tag{19.2}$$

As will be discussed later, VaR has a serious deficiency—it discourages diversification—and for this reason it is being replaced by newer risk measures. One of these newer risk measures is the expected loss given that the loss exceeds VaR, which is called by a variety of names: *expected shortfall*, the *expected loss given a tail event*, *tail loss*, and *shortfall*. The name *expected shortfall* and the abbreviation ES will be used here.

For any loss distribution, continuous or not,

$$\text{ES}(\alpha) = \frac{\int_0^\alpha \text{VaR}(u)\, du}{\alpha}, \tag{19.3}$$

which is the average of $\text{VaR}(u)$ over all u that are less than or equal to α. If $\mathcal{L}$ has a continuous distribution,

$$\text{ES}(\alpha) = E\left\{\mathcal{L}\middle|\mathcal{L} > \text{VaR}(\alpha)\right\} = E\left\{\mathcal{L}\middle|\mathcal{L} \geq \text{VaR}(\alpha)\right\}. \tag{19.4}$$

Example 19.1. VaR with a normally distributed loss

Suppose that the yearly return on a stock is normally distributed with mean 0.04 and standard deviation 0.18. If one purchases \$100,000 worth of this stock, what is the VaR with T equal to one year?

To answer this question, we use the fact that the loss distribution is normal with mean -4000 and standard deviation 18,000, with all units in dollars. Therefore, VaR is

$$-4000 + 18{,}000 z_\alpha,$$

where z_α is the α-upper quantile of the standard normal distribution. $\text{VaR}(\alpha)$ is plotted as a function of α in Figure 19.1. VaR depends heavily on α and in this figure ranges from 46,527 when α is 0.025 to 8,226 when α is 0.25.

□

In applications, risk measures will rarely, if ever, be known exactly as in these simple examples. Instead, risk measures are estimated, and estimation error is another source of uncertainty. This uncertainty can be quantified using a confidence interval for the risk measure. We turn next to these topics.

19.2 Estimating VaR and ES with One Asset

To illustrate the techniques for estimating VaR and ES, we begin with the simple case of a single asset. In this section, these risk measures are estimated using historic data to estimate the distribution of returns. We make the assumption that returns are stationary, at least over the historic period we use.

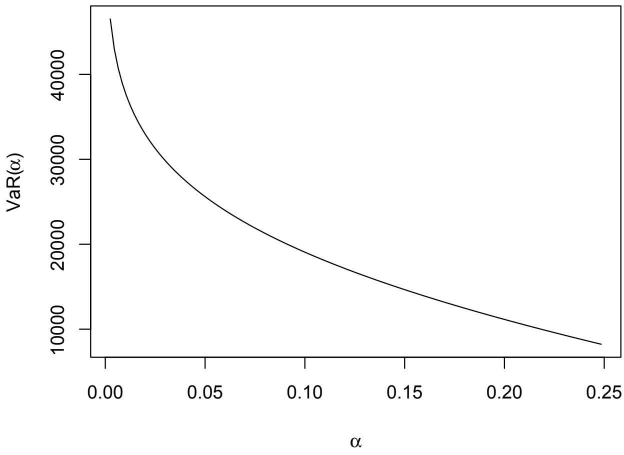

Fig. 19.1. VaR(α) *for* $0.025 < \alpha < 0.25$ *when the loss distribution is normally distributed with mean* -4000 *and standard deviation* 18,000*.*

This is usually a reasonable assumption. We will also assume that the returns are independent. Independence is a much less reasonable assumption because of volatility clustering, and later we will remove this assumption by using GARCH models.

Two cases are considered, first without and then with the assumption of a parametric model for the return distribution.

19.2.1 Nonparametric Estimation of VaR and ES

We start with *nonparametric* estimates of VaR and ES, meaning that the loss distribution is not assumed to be in a parametric family such as the normal or t-distributions.

Suppose that we want a confidence coefficient of $1-\alpha$ for the risk measures. Therefore, we estimate the α-quantile of the return distribution, which is the α-upper quantile of the loss distribution. In the nonparametric method, this quantile is estimated as the α-quantile of a sample of historic returns, which we will call $\widehat{q}(\alpha)$. If S is the size of the current position, then the nonparametric estimate of VaR is

$$\widehat{\text{VaR}}^{\text{np}}(\alpha) = -S \times \widehat{q}(\alpha),$$

with the minus sign converting revenue (return times initial investment) to a loss. In this chapter, superscripts and subscripts will sometimes be placed on VaR and ES to provide information. Here, the superscript "np" means "nonparametrically estimated."

To estimate ES, let $R_1, \ldots, R_n$ be the historic returns and define $\mathcal{L}_i = -S \times R_i$. Then

$$\widehat{\mathrm{ES}}^{\mathrm{np}}(\alpha) = \frac{\sum_{i=1}^n \mathcal{L}_i \, I\{\mathcal{L}_i > \widehat{\mathrm{VaR}}(\alpha)\}}{\sum_{i=1}^n I\{\mathcal{L}_i > \widehat{\mathrm{VaR}}(\alpha)\}} = -S \times \frac{\sum_{i=1}^n R_i \, I\{R_i < \widehat{q}(\alpha)\}}{\sum_{i=1}^n I\{R_i < \widehat{q}(\alpha)\}}, \quad (19.5)$$

which is the average of all $\mathcal{L}_i$ exceeding $\widehat{\mathrm{VaR}}^{\mathrm{np}}(\alpha)$. Here $I\{\mathcal{L}_i > \widehat{\mathrm{VaR}}^{\mathrm{np}}(\alpha)\}$ is the indicator that $\mathcal{L}_i$ exceeds $\widehat{\mathrm{VaR}}^{\mathrm{np}}(\alpha)$ and similarly for $I\{R_i < \widehat{q}(\alpha)\}$.

Example 19.2. Nonparametric VaR and ES for a position in an S&P 500 index fund

As a simple example, suppose that you hold a \$20,000 position in an S&P 500 index fund, so your returns are those of this index, and that you want a 24-hour VaR. We estimate this VaR using the 1000 daily returns on the S&P 500 for the period ending in April 1991. These log returns are a subset of the data set `SP500` in `R`'s `Ecdat` package. The full time series is plotted in Figure 4.1. Black Monday, with a log return of −0.23, occurs near the beginning of the shortened time series used in this example.

Suppose you want 95% confidence. The 0.05 quantile of the returns computed by `R`'s `quantile` function is −0.0169. In other words, a daily return of −0.0169 or less occurred only 5% of the time in the historic data, so we estimate that there is a 5% chance of a return of that size occurring during the next 24 hours. A return of −0.0169 on a \$20,000 investment yields a revenue of −\$337.43, and therefore the estimated $\widehat{\mathrm{VaR}}(0.05, \text{24 hours})$ is \$337.43.

ES(0.05) is obtained by averaging all returns below −0.0169 and multiplying this average by −20,000. The result is $\widehat{\mathrm{ES}}^{\mathrm{np}}(0.05) = \$619.3$.

□

19.2.2 Parametric Estimation of VaR and ES

Parametric estimation of VaR and ES has a number of advantages. For example, parametric estimation allows the use of GARCH models to adapt the risk measures to the current estimate of volatility. Also, risk measures can be easily computed for a portfolio of stocks if we assume that their returns have a joint parametric distribution such as a multivariate t-distribution. Nonparametric estimation using sample quantiles works best when the sample size and α are reasonably large. With smaller sample sizes or smaller values of α, it is preferable to use parametric estimation. In this section, we look at parametric estimation of VaR and ES when there is a single asset.

Let $F(y|\boldsymbol{\theta})$ be a parametric family of distributions used to model the return distribution and suppose that $\widehat{\boldsymbol{\theta}}$ is an estimate of $\boldsymbol{\theta}$, such as, the MLE computed from historic returns. Then $F^{-1}(\alpha|\widehat{\boldsymbol{\theta}})$ is an estimate of the α-quantile of the return distribution and

$$\widehat{\mathrm{VaR}}^{\mathrm{par}}(\alpha) = -S \times F^{-1}(\alpha|\widehat{\boldsymbol{\theta}}) \tag{19.6}$$

is a parametric estimate of VaR(α). As before, S is the size of the current position.

Let $f(y|\boldsymbol{\theta})$ be the density of $F(y|\boldsymbol{\theta})$. Then the estimate of expected shortfall is

$$\widehat{\mathrm{ES}}^{\mathrm{par}}(\alpha) = -\frac{S}{\alpha} \times \int_{-\infty}^{F^{-1}(\alpha|\boldsymbol{\theta})} x f(x|\widehat{\boldsymbol{\theta}})\, dx. \tag{19.7}$$

The superscript "par" denotes "parametrically estimated." Computing this integral is not always easy, but in the important cases of normal and t-distributions there are convenient formulas.

Suppose the return has a t-distribution with mean equal to μ, scale parameter equal to λ, and ν degrees of freedom. Let f_ν and F_ν be, respectively, the t-density and t-distribution function with ν degrees of freedom. The expected shortfall is

$$\widehat{\mathrm{ES}}^{\mathrm{t}}(\alpha) = S \times \left\{ -\mu + \lambda \left(\frac{f_\nu\{F_\nu^{-1}(\alpha)\}}{\alpha} \left[\frac{\nu + \{F_\nu^{-1}(\alpha)\}^2}{\nu - 1} \right] \right) \right\}. \tag{19.8}$$

The formula for normal loss distributions is obtained by a direct calculation or letting $\nu \to \infty$ in (19.8). The result is

$$\mathrm{ES}^{\mathrm{norm}}(\alpha) = S \times \left\{ -\mu + \sigma \left(\frac{\phi\{\Phi^{-1}(\alpha)\}}{\alpha} \right) \right\}, \tag{19.9}$$

where μ and σ are the mean and standard deviation of the returns and ϕ and Φ are the standard normal density and CDF. The superscripts "t" and "norm" denote estimates assuming a normal return and t-distributed return, respectively.

Parametric estimation with one asset is illustrated in the next example.

Example 19.3. Parametric VaR and ES for a position in an S&P 500 index fund

This example uses the same data set as in Example 19.2 so that parametric and nonparametric estimates can be compared. We will assume that the returns are i.i.d. with a t-distribution. Under this assumption, VaR is

$$\widehat{\mathrm{VaR}}^{\mathrm{t}}(\alpha) = -S \times \{\widehat{\mu} + q_{\alpha,t}(\widehat{\nu})\widehat{\lambda}\}, \tag{19.10}$$

where $\widehat{\mu}$, $\widehat{\lambda}$, and $\widehat{\nu}$ are the estimated mean, scale parameter, and degrees of freedom of a sample of returns. Also, $q_{\alpha,t}(\widehat{\nu})$ is the α-quantile of the t-distribution with $\widehat{\nu}$ degrees of freedom, so that $\{\widehat{\mu} + q_{\alpha,t}(\widehat{\nu})\widehat{\lambda}\}$ is the αth quantile of the fitted distribution.

The t-distribution was fit using R's `fitdistr` function and the estimates were $\widehat{\mu} = 0.000689$, $\widehat{\lambda} = 0.007164$, and $\widehat{\nu} = 2.984$. For later reference, the estimated standard deviation is $\widehat{\sigma} = \widehat{\lambda}\sqrt{\widehat{\nu}/(\widehat{\nu}-2)} = 0.01248$.

The 0.05-quantile of the t-distribution with 2.984 degrees of freedom is -2.3586. Therefore, by (19.6),

$$\widehat{\mathrm{VaR}}^{\mathrm{t}}(0.05) = -20000 \times \{0.000689 - (2.3586)(0.007164)\} = \$323.42.$$

Notice that the nonparametric estimate, $\widehat{\mathrm{VaR}}^{\mathrm{np}}(0.05) = \$337.55$, is similar to but somewhat larger than the parametric estimate, \$323.42.

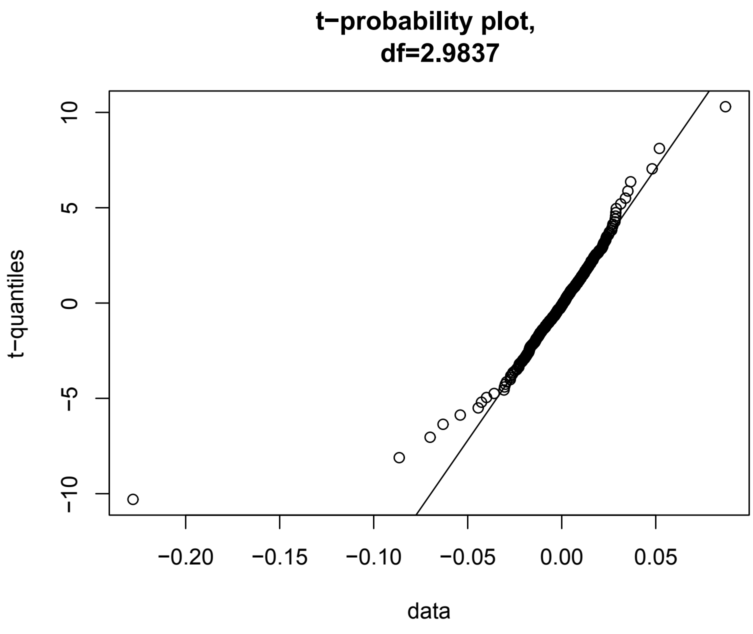

Fig. 19.2. *t-plot of the S&P 500 returns used in Examples 19.2 and 19.3. The deviations from linearity in the tails, especially the left tail, indicate that the t-distribution does not fit the data in the extreme tails. The reference line goes through the first and third quartiles. The t-quantiles use 2.9837 degrees of freedom, the MLE.*

The parametric estimate of $\mathrm{ES}^t(0.05)$ is \$543.81 and is found by substituting $S = 20{,}000$, $\alpha = 0.05$, $\widehat{\mu} = 0.000689$, $\widehat{\lambda} = 0.007164$, and $\widehat{\nu} = 2.984$ into (19.8). The parametric estimate of $\mathrm{ES}^t(0.05)$ is noticeably shorter than the nonparametric. The reason the two estimates differ is that the extreme left tail of the returns, roughly the smallest 10 of 1000 returns, is heavier than the tail of a t-distribution with 2.984 degrees of freedom; see the t-plot in Figure 19.2.

□

19.3 Confidence Intervals for VaR and ES Using the Bootstrap

The estimates of VaR and ES are precisely that, just estimates. If we had used a different sample of historic data, then we would have gotten different estimates of these risk measures. We just calculated VaR and ES values to five significant digits, but do we really have that much precision? The reader has probably guessed (correctly) that we do not, but how much precision do we have? How can we learn the true precision of the estimates? Fortunately, a confidence interval for VaR or ES is rather easily obtained by bootstrapping. Any of the confidence interval procedures in Section 6.3 can be used. We will see that even with 1000 returns to estimate VaR and ES, these risk measures are estimated with considerable uncertainty.

For now, we will assume an i.i.d. sample of historic returns and use model-free resampling. In Section 19.4 we will allow for dependencies, for instance, GARCH effects, in the data and we will use model-based resampling.

Suppose we have a large number, B, of resamples of the returns data. Then a $\mathrm{VaR}(\alpha)$ or $\mathrm{ES}(\alpha)$ estimate is computed from each resample and for the original sample. The confidence interval can be based upon either a parametric or nonparametric estimator of $\mathrm{VaR}(\alpha)$ or $\mathrm{ES}(\alpha)$. Suppose that we want the confidence coefficient of the interval to be $1-\gamma$. The interval's confidence coefficient should not be confused with the confidence coefficient of VaR, which we denote by $1-\alpha$. The $\gamma/2$-lower and -upper quantiles of the bootstrap estimates of $\mathrm{VaR}(\alpha)$ and $\mathrm{ES}(\alpha)$ are the limits of the basic percentile method confidence intervals.

It is worthwhile to restate the meanings of α and γ, since it is easy to confuse these two confidence coefficients, but they need to be distinguished since they have rather different interpretations. $\mathrm{VaR}(\alpha)$ is defined so that the probability of a loss being greater than $\mathrm{VaR}(\alpha)$ is α. On the other hand, γ is the confidence coefficient for the confidence interval for $\mathrm{VaR}(\alpha)$ and $\mathrm{ES}(\alpha)$. If many confidence intervals are constructed, then approximately γ of them do not contain the true risk measure. Thus, α is about the loss from the investment while γ is about the confidence interval being correct. An alternative way to view the difference between α and γ is that $\mathrm{VaR}(\alpha)$ and $\mathrm{ES}(\alpha)$ are measuring risk due to uncertainty about future losses, assuming perfect knowledge of the loss distribution, while the confidence intervals tell us the uncertainty of these risk measures due to imperfect knowledge of the loss distribution.

Example 19.4. Bootstrap confidence intervals for VaR and ES for a position in an S&P 500 index fund

In this example, we continue Examples 19.2 and 19.3 and find a confidence interval for VaR(α) and ES(α). We use $\alpha = 0.05$ as before and $\gamma = 0.1$. $B =$ 5,000 resamples were taken.

The basic percentile confidence intervals for VaR(0.05) were (297, 352) and (301, 346) using nonparametric and parametric estimators of VaR(0.05), respectively. For ES(0.05), the corresponding basic percentile confidence intervals were (487, 803) and (433, 605). We see that there is considerable uncertainty in the risk measures, especially for ES(0.05) and especially using nonparametric estimation.

The bootstrap computation took 33.3 minutes using an `R` program and a 2.13 GHz Pentium™ processor running under Windows™. The computations took this long because the optimization step to find the MLE for parametric estimation is moderately expensive in computational time, at least if it is repeated 5000 times.

Waiting over a half an hour for the confidence interval may not be an attractive proposition. However, a reasonable measure of precision can be obtained with far fewer bootstrap repetitions. One might use only 50 repetitions, which would take less than a minute. This is not enough resamples to use basic percentile bootstrap confidence intervals, but instead one can use the normal approximation bootstrap confidence interval, (6.4). As an example, the normal approximation interval for the nonparametric estimate of VaR(0.05) is (301, 361) using only the first 50 bootstrap resamples. This interval gives the same general impression of accuracy as the above basic percentile method interval, (297, 352), that uses all 5000 resamples.

The normal approximation interval assumes that $\widehat{\text{VaR}}(0.05)$ is approximately normally distributed. This assumption is justified by the central limit theorem for sample quantiles (Section 4.3.1) and the fact that $\widehat{\text{VaR}}(0.05)$ is a multiple of a sample quantile. The normal approximation does *not* require that the returns are normally distributed. In fact, we are modeling them as t-distributed when computing the parametric estimates.

□

19.4 Estimating VaR and ES Using ARMA/GARCH Models

As we have seen in Chapters 9 and 18, daily equity returns typically have a small amount of autocorrelation and a greater amount of volatility clustering.

When calculating risk measures, the autocorrelation can be ignored if it is small enough, but the volatility clustering is less ignorable. In this section, we use ARMA/GARCH models so that VaR(α) and ES(α) can adjust to periods of high or low volatility.

Assume that we have n returns, $R_1, \ldots, R_n$ and we need to estimate VaR and ES for the next return R_{n+1}. Let $\widehat{\mu}_{n+1|n}$ and $\widehat{\sigma}_{n+1|n}$ be the estimated conditional mean and variance of tomorrow's return R_{n+1} conditional on the current information set, which in this context is simply $\{R_1, \ldots, R_n\}$. We will also assume that R_{n+1} has a conditional t-distribution with ν degrees of freedom. After fitting an ARMA/GARCH model, we have estimates of $\widehat{\nu}$, $\widehat{\mu}_{n+1|n}$, and $\widehat{\sigma}_{n+1|n}$. The estimated conditional scale parameter is

$$\widehat{\lambda}_{n+1|n} = \sqrt{(\widehat{\nu}-2)/\widehat{\nu}}\ \widehat{\sigma}_{n+1|n}. \tag{19.11}$$

VaR and ES are estimated as in Section 19.2.2 but with $\widehat{\mu}$ and $\widehat{\lambda}$ replaced by $\widehat{\mu}_{n+1|n}$ and $\widehat{\lambda}_{n+1|n}$.

Example 19.5. VaR and ES for a position in an S&P 500 index fund using a GARCH(1,1) model

An AR(1)/GARCH(1,1) model was fit to the log returns on the S&P 500. The AR(1) coefficient was small and not significantly different from 0, so a GARCH(1,1) was used for estimation of VaR and ES. The GARCH(1,1) fit is

```
Call: garchFit(formula = ~garch(1, 1), data = SPreturn,
   cond.dist = "std")

Error Analysis:
        Estimate  Std. Error  t value Pr(>|t|)
mu      7.147e-04   2.643e-04    2.704  0.00685 **
omega   2.833e-06   9.820e-07    2.885  0.00392 **
alpha1  3.287e-02   1.164e-02    2.824  0.00474 **
beta1   9.384e-01   1.628e-02   57.633  < 2e-16 ***
shape   4.406e+00   6.072e-01    7.256    4e-13 ***
```

The conditional mean and standard deviation of the next return were estimated to be 0.00071 and 0.00950. For the estimation of VaR and ES, the next return was assumed to have a t-distribution with these values for the mean and standard deviation and 4.406 degrees of freedom. The estimate of VaR was \$277.21 and the estimate of ES was \$414.61. The VaR and ES estimates using the GARCH model are considerably smaller than the parametric estimates in Example 19.2 (\$323.42 and \$543.81), because the conditional standard deviation used here (0.00950) is smaller than the marginal standard deviation (0.01248) used in Example 19.2; see Figure 19.3, where the dashed horizontal line's height is the marginal standard deviation and the conditional

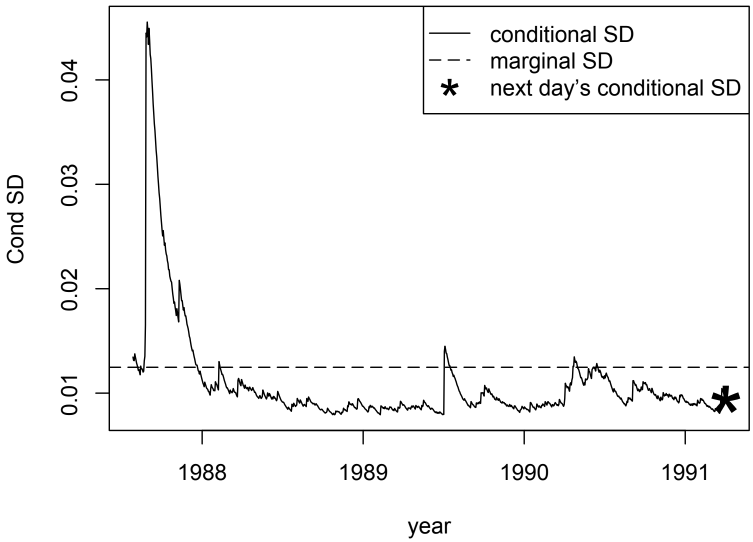

Fig. 19.3. *Conditional standard deviation of the S&P 500 returns based on a GARCH(1,1) model. The asterisk is at the conditional standard deviation of the next day's return after the end of the series, and the height of the horizontal line is the marginal standard deviation.*

standard deviation of the next day's return is indicated by a large asterisk. The marginal standard deviation is inflated by periods of higher volatility such as in October 1987 (near Black Monday) on the left-hand side of Figure 19.3. □

19.5 Estimating VaR and ES for a Portfolio of Assets

When VaR is estimated for a portfolio of assets rather than a single asset, parametric estimation based on the assumption of multivariate normal or t-distributed returns is very convenient, because the portfolio's return will have a univariate normal or t-distributed return. The portfolio theory and factor models developed in Chapters 11 and 17 can be used to estimate the mean and variance of the portfolio's return.

Estimating VaR becomes complex when the portfolio contains stocks, bonds, options, foreign exchange positions, and other assets. However, when a portfolio contains only stocks, then VaR is relatively straightforward to estimate, and we will restrict attention to this case—see Section 19.10 for discussion of the literature covering more complex cases.

With a portfolio of stocks, means, variances, and covariances of returns could be estimated directly from a sample of returns as discussed in Chapter 11 or using a factor model as discussed in Section 17.4.2. Once these estimates are available, they can be plugged into equations (11.6) and (11.7) to obtain estimates of the expected value and variance of the return on the portfolio, which are denoted by $\widehat{\mu}_P$ and $\widehat{\sigma}_P^2$. Then, analogous to (19.10), VaR can be estimated, assuming normally distributed returns on the portfolio (denoted with a subscript "P"), by

$$\widehat{\mathrm{VaR}}_P^{\mathrm{norm}}(\alpha) = -S \times \{\widehat{\mu}_P + \Phi^{-1}(\alpha)\widehat{\sigma}_P\}, \tag{19.12}$$

where S is the initial value of the portfolio. Moreover, using (19.9), the estimated expected shortfall is

$$\widehat{\mathrm{ES}}_P^{\mathrm{norm}}(\alpha) = S \times \left\{-\widehat{\mu}_P + \widehat{\sigma}_P \left(\frac{\phi\{\Phi^{-1}(\alpha)\}}{\alpha}\right)\right\}. \tag{19.13}$$

If the stock returns have a joint t-distribution, then the returns on the portfolio have a univariate t-distribution with the same degrees of freedom, and VaR and ES for the portfolio can be calculated using formulas in Section 19.2.2. If the returns on the portfolio have a t-distribution with mean μ_P, scale parameter λ_P, and degrees of freedom ν, then the estimated VaR is

$$\widehat{\mathrm{VaR}}_P^{\mathrm{t}}(\alpha) = -S\{\widehat{\mu}_P + F_\nu^{-1}(\alpha)\widehat{\lambda}_P\}, \tag{19.14}$$

and the estimated expected shortfall is

$$\widehat{\mathrm{ES}}_P^{\mathrm{t}}(\alpha) = S \times \left\{-\widehat{\mu}_P + \widehat{\lambda}_P \left(\frac{f_{\widehat{\nu}}\{F_{\widehat{\nu}}^{-1}(\alpha)\}}{\alpha}\left[\frac{\widehat{\nu} + \{F_{\widehat{\nu}}^{-1}(\alpha)\}^2}{\widehat{\nu}-1}\right]\right)\right\}. \tag{19.15}$$

Example 19.6. VaR and ES for portfolios of the three stocks in the `CRSPday` *data set*

This example uses the data set `CRSPday` used earlier in Examples 7.1 and 7.4. There are four variables—returns on GE, IBM, Mobil, and the CRSP index and we found in Example 7.4 that their returns can be modeled as having a multivariate t-distribution with 5.94 degrees of freedom. In this example, we will only the returns on the three stocks. The t-distribution parameters were reestimated without the CRSP index and $\widehat{\nu}$ changed slightly to 5.81.

The estimated mean was

$$\widehat{\mu} = (\,0.0008584 \quad 0.0003249 \quad 0.0006162\,)^{\mathsf{T}}$$

and the estimated covariance matrix was

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 1.273\text{e}-04 & 5.039\text{e}-05 & 3.565\text{e}-05 \\ 5.039\text{e}-05 & 1.812\text{e}-04 & 2.400\text{e}-05 \\ 3.565\text{e}-05 & 2.400\text{e}-05 & 1.149\text{e}-04 \end{pmatrix}.$$

For an equally weighted portfolio with $w = (\,1/3 \quad 1/3 \quad 1/3\,)^{\mathsf{T}}$, the mean return for the portfolio is estimated to be

$$\widehat{\mu}_P = \widehat{\mu}^{\mathsf{T}} w = 0.0005998$$

and the standard deviation of the portfolio's return is estimated as

$$\widehat{\sigma}_P = \sqrt{w^{\mathsf{T}} \widehat{\boldsymbol{\Sigma}} w} = 0.008455.$$

The return on the portfolio has a t-distribution with this mean and standard deviation and the same degrees of freedom as the multivariate t-distribution of the three stock returns. The scale parameter, using $\widehat{\nu} = 5.81$, is

$$\widehat{\lambda}_P = \sqrt{(\widehat{\nu} - 2)/\widehat{\nu}} \times 0.008455 = 0.006847.$$

Therefore,

$$\widehat{\text{VaR}}^{t}(0.05) = -S\left\{\widehat{\mu}_P + \widehat{\lambda}_P\, \widehat{q}_{0.05,t}(\widehat{\nu})\right\} = S \times 0.01278,$$

so, for example, with $S = \$20{,}000$, $\widehat{\text{VaR}}^{t}(0.05) = \$256$.

The estimated ES using (19.8) and $S = \$20{,}000$ is

$$\widehat{\text{ES}}^{t}(0.05) = S \times \left\{ -\widehat{\mu}_P + \widehat{\lambda}_P \left(\frac{f_{\widehat{\nu}}\{\widehat{q}_{0.05,t}(\widehat{\nu})\}}{\alpha} \left[\frac{\widehat{\nu} + \{\widehat{q}_{0.05,t}(\widehat{\nu})\}^2}{\widehat{\nu} - 1} \right] \right) \right\} = \$363.$$

□

19.6 Estimation of VaR Assuming Polynomial Tails

There is an interesting compromise between using a totally nonparametric estimator of VaR as in Section 19.2.1 and a parametric estimator as in Section 19.2.2. The nonparametric estimator is feasible for large α, but not for small α. For example, if the sample had 1000 returns, then reasonably accurate estimation of the 0.05-quantile is feasible, but not estimation of the 0.0005-quantile. Parametric estimation can estimate VaR for any value of α but is sensitive to misspecification of the tail when α is small. Therefore, a methodology intermediary between totally nonparametric and parametric estimation is attractive.

The approach used in this section assumes that the return density has a polynomial left tail, or equivalently that the loss density has a polynomial right tail. Under this assumption, it is possible to use a nonparametric estimate of $\mathrm{VaR}(\alpha_0)$ for a *large* value of α_0 to obtain estimates of $\mathrm{VaR}(\alpha_1)$ for *small* values of α_1. It is assumed here that $\mathrm{VaR}(\alpha_1)$ and $\mathrm{VaR}(\alpha_0)$ have the same horizon T.

Because the return density is assumed to have a polynomial left tail, the return density f satisfies

$$f(y) \sim Ay^{-(a+1)}, \text{ as } y \to -\infty, \tag{19.16}$$

where $A > 0$ is a constant and $a > 0$ is the tail index. Therefore,

$$P(R \le y) \sim \int_{-\infty}^{y} f(u)\,du = \frac{A}{a} y^{-a}, \text{ as } y \to -\infty, \tag{19.17}$$

and if $y_1 > 0$ and $y_2 > 0$, then

$$\frac{P(R < -y_1)}{P(R < -y_2)} \approx \left(\frac{y_1}{y_2}\right)^{-a}. \tag{19.18}$$

Now suppose that $y_1 = \mathrm{VaR}(\alpha_1)$ and $y_2 = \mathrm{VaR}(\alpha_0)$, where $0 < \alpha_1 < \alpha_0$. Then (19.18) becomes

$$\frac{\alpha_1}{\alpha_0} = \frac{P\{R < -\mathrm{VaR}(\alpha_1)\}}{P\{R < -\mathrm{VaR}(\alpha_0)\}} \approx \left(\frac{\mathrm{VaR}(\alpha_1)}{\mathrm{VaR}(\alpha_0)}\right)^{-a} \tag{19.19}$$

or

$$\frac{\mathrm{VaR}(\alpha_1)}{\mathrm{VaR}(\alpha_0)} \approx \left(\frac{\alpha_0}{\alpha_1}\right)^{1/a},$$

so, now dropping the subscript “1” of α_1 and writing the approximate equality as exact, we have

$$\mathrm{VaR}(\alpha) = \mathrm{VaR}(\alpha_0)\left(\frac{\alpha_0}{\alpha}\right)^{1/a}. \tag{19.20}$$

Equation (19.20) becomes an estimate of $\mathrm{VaR}(\alpha)$ when $\mathrm{VaR}(\alpha_0)$ is replaced by a nonparametric estimate and the tail index a is replaced by one of the estimates discussed soon in Section 19.6.1. Notice another advantage of (19.20), that it provides an estimate of $\mathrm{VaR}(\alpha)$ not just for a single value of α but for all values. This is useful if one wants to compute and compare $\mathrm{VaR}(\alpha)$ for a variety of values of α, as is illustrated in Example 19.7 ahead. The value of α_0 must be large enough that $\mathrm{VaR}(\alpha_0)$ can be accurately estimated, but α can be any value less than α_0.

A model combining parametric and nonparametric components is called *semiparametric*, so estimator (19.20) is semiparametric because the tail index is specified by a parameter, but otherwise the distribution is unspecified.

To find a formula for ES, we will assume further that for some $c < 0$, the returns density satisfies

$$f(y) = A|y|^{-(a+1)}, \quad y \le c, \tag{19.21}$$

so that we have equality in (19.16) for $y \le c$. Then, for any $d \le c$,

$$P(R \le d) = \int_{-\infty}^{d} A|y|^{-(a+1)}\, dy = \frac{A}{a}|d|^{-a}, \tag{19.22}$$

and the conditional density of R given that $R \le d$ is

$$f(y|R \le d) = \frac{Ay^{-(a+1)}}{P(R \le d)} = a|d|^{a}|y|^{-(a+1)}. \tag{19.23}$$

It follows from (19.23) that for $a > 1$,

$$E\Big(|R| \,\Big|\, R \le d\Big) = a|d|^{a} \int_{-\infty}^{d} |y|^{-a} dy = \frac{a}{a-1}|d|. \tag{19.24}$$

(For $a \le 1$, this expectation is $+\infty$.) If we let $d = -\mathrm{VaR}(\alpha)$, then we see that

$$\mathrm{ES}(\alpha) = \frac{a}{a-1}\mathrm{VaR}(\alpha) = \frac{1}{1-a^{-1}}\mathrm{VaR}(\alpha), \text{ if } a > 1. \tag{19.25}$$

Formula (19.25) enables one to estimate $\mathrm{ES}(\alpha)$ using an estimate of $\mathrm{VaR}(\alpha)$ and an estimate of a.

19.6.1 Estimating the Tail Index

In this section, we estimate the tail index assuming a polynomial left tail. Two estimators will be introduced, the regression estimator and the Hill estimator.

Regression Estimator of the Tail Index

It follows from (19.17) that

$$\log\{P(R \le -y)\} = \log(L) - a\log(y), \tag{19.26}$$

where $L = A/a$.

If $R_{(1)}, \ldots, R_{(n)}$ are the order statistics of the returns, then the number of observed returns less than or equal to $R_{(k)}$ is k, so we estimate $\log\{P(R \le R_{(k)})\}$ to be $\log(k/n)$. Then, from (19.26), we have

$$\log(k/n) \approx \log(L) - a\log(-R_{(k)}) \tag{19.27}$$

or, rearranging (19.27),

$$\log(-R_{(k)}) \approx (1/a)\log(L) - (1/a)\log(k/n). \tag{19.28}$$

The approximation (19.28) is expected to be accurate only if $-R_{(k)}$ is large, which means k is small, perhaps only 5%, 10%, or 20% of the sample size n. If we plot the points $\big[\{\log(k/n), \log(-R_{(k)})\}\big]_{k=1}^{m}$ for m equal to a small percentage of n, say 10%, then we should see these points fall on roughly a straight line. Moreover, if we fit the straight-line model (19.28) to these points by least squares, then the estimated slope, call it $\widehat{\beta}_1$, estimates $-1/a$. Therefore, we will call $-1/\widehat{\beta}_1$ the *regression estimator of the tail index.*

Hill Estimator

The Hill estimator of the left tail index a of the return density f uses all data less than a constant c, where c is sufficiently small that

$$f(y) = A|y|^{-(a+1)} \tag{19.29}$$

is assumed to be true for $y < c$. The choice of c is crucial and will be discussed below. Let $Y_{(1)}, \ldots, Y_{(n)}$ be order statistics of the returns and $n(c)$ be the number of Y_1 less than or equal to c. By (19.23), the conditional density of Y_i given that $Y_i \le c$ is

$$a|c|^a|y|^{-(a+1)}. \tag{19.30}$$

Therefore, the likelihood for $Y_{(1)}, \ldots, Y_{(n(c))}$ is

$$L(a) = \left(\frac{a|c|^a}{|Y_1|^{a+1}}\right)\left(\frac{a|c|^a}{|Y_2|^{a+1}}\right)\cdots\left(\frac{a|c|^a}{|Y_{n(c)}|^{a+1}}\right),$$

and the log-likelihood is

$$\log\{L(a)\} = \sum_{i=1}^{n(c)}\{\log(a) + a\log(|c|) - (a+1)\log(|Y_{(i)}|)\}. \tag{19.31}$$

Differentiating the right-hand side of (19.31) with respect to a and setting the derivative equal to 0 gives the equation

$$\frac{n(c)}{a} = \sum_{i=1}^{n(c)} \log\left(Y_{(i)}/c\right).$$

Therefore, the MLE of a, which is called the *Hill estimator*, is

$$\widehat{a}^{\text{Hill}}(c) = \frac{n(c)}{\sum_{i=1}^{n(c)} \log\left(Y_{(i)}/c\right)}. \tag{19.32}$$

Note that $Y_{(i)} \le c < 0$, so that $Y_{(i)}/c$ is positive.

How should c be chosen? Usually c is equal to one of $Y_1, \ldots, Y_n$ so that $c = Y_{(n(c))}$, and therefore choosing c means choosing $n(c)$. The choice involves a bias–variance tradeoff. If $n(c)$ is too large, then $f(y) = A|y|^{-(a+1)}$ will not hold for all values of $y \le c$, causing bias. If $n(c)$ is too small, then there will be too few Y_i below c and $\widehat{a}^{\text{Hill}}(c)$ will be highly variable and unstable because it uses too few data. However, we can hope that there is a range of values of $n(c)$ where $\widehat{a}^{\text{Hill}}(c)$ is reasonably constant because it is neither too biased nor too variable.

A *Hill plot* is a plot of $\widehat{a}^{\text{Hill}}(c)$ versus $n(c)$ and is used to find this range of values of $n(c)$. In a Hill plot, one looks for a range of $n(c)$ where the estimator is nearly constant and then chooses $n(c)$ in this range.

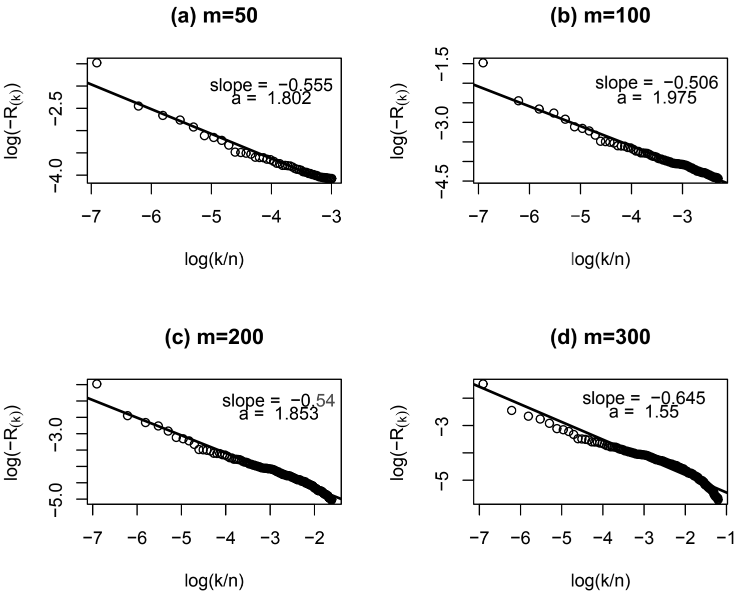

Fig. 19.4. *Plots for estimating the left tail index of the S&P* 500 *returns by regression. "Slope" is the least-squares slope estimate and "a" is* $-1/$*slope.*

Example 19.7. Estimating the left tail index of the S&P 500 returns

This example uses the 1000 daily S&P 500 returns used in Examples 19.2 and 19.3. First, the regression estimator of the tail index was calculated. The values $\left[\{\log(k/n), \log(-R_{(k)})\}\right]_{k=1}^{m}$ were plotted for $m = 50$, 100, 200, and 300 to find the largest value of m giving a roughly linear plot and $m = 100$ was selected. The plotted points and the least-squares lines can be seen in Figure 19.4. The slope of the line with $m = 100$ was -0.506, so a was estimated to be $1/0.506 = 1.975$.

Suppose we have invested \$20,000 in an S&P 500 index fund. We will use $\alpha_0 = 0.1$. VaR(0.1, 24 hours) is estimated to be $-$\$20,000 times the 0.1-quantile of the 1000 returns. The sample quantile is -0.0117, so $\widehat{\mathrm{VaR}}^{\mathrm{np}}(0.1,$ 24 hours) = \$234. Using (19.20) and $a = 1.975$ ($1/a = 0.506$), we have

$$\widehat{\mathrm{VaR}}(\alpha) = 234\left(\frac{0.1}{\alpha}\right)^{0.506}. \tag{19.33}$$

The solid curve in Figure 19.5 is a plot of $\widehat{\mathrm{VaR}}(\alpha)$ for $0.0025 \le \alpha \le 0.25$ using (19.33) and the regression estimator of a. The curve with short dashes is the same plot but with the Hill estimator of a, which is 2.2—see below. The

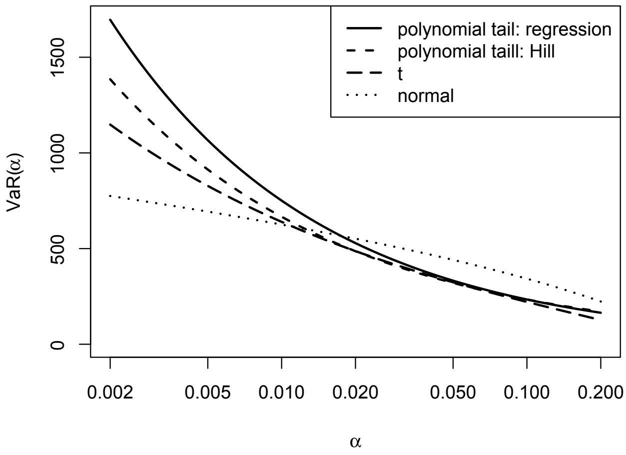

Fig. 19.5. *Estimation of VaR(α) using formula* (19.33) *and the regression estimator of the tail index (solid), using formula* (19.33) *and the Hill estimator of the tail index (short dashes), assuming t-distributed returns (long dashes), and assuming normally distributed returns (dotted). Note the log-scale on the x-axis.*

curve with long dashes is $\text{VaR}(\alpha)$ estimated assuming t-distributed returns as discussed in Section 19.2.2, and the dotted curve is estimated assuming normally distributed returns. The return distribution has much heavier tails than a normal distribution, and the latter curve is included only to show the effect of model misspecification. The parametric estimates based on the t-distribution are similar to the estimates assuming a polynomial tail except when α is very small. The difference between the two estimates for small α ($\alpha < 0.01$) is to be expected because the polynomial tail with tail index 1.975 or 2.2 is heavier than the tail of the t-distribution with $\nu = a = 2.984$. If α is in the range 0.01 to 0.2, then $\widehat{\text{VaR}}(\alpha)$ is relatively insensitive to the choice of model, except for the poorly fitting normal model. This is a good reason for preferring $\alpha \geq 0.01$.

It follows from (19.25) using the regression estimate $\widehat{a} = 1.975$ that

$$\widehat{\text{ES}}(\alpha) = \frac{1.975}{0.975}\,\widehat{\text{VaR}}(\alpha) = 2.026\,\widehat{\text{VaR}}(\alpha). \tag{19.34}$$

The Hill estimator of a was also implemented. Figure 19.6 contains Hill plots, that is, plots of the Hill estimate $\widehat{a}_{\text{Hill}}(c)$ versus $n(c)$. In panel (a), $n(c)$ ranges from 25 to 250. There seems to be a region of stability when $n(c)$ is between 25 and 120, which is shown in panel (b). In panel (b), we see a region of even greater stability when $n(c)$ is between 60 and 100. Panel (c) zooms in

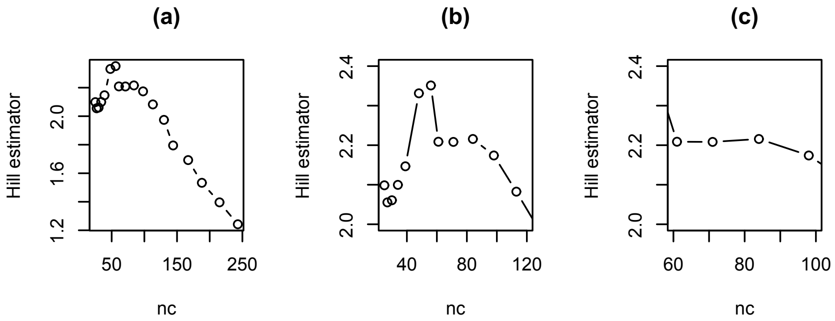

Fig. 19.6. *Estimation of tail index by applying a Hill plot to the daily returns on the S&P* 500 *for* 1000 *consecutive trading days ending on March* 4, 2003*. (a) Full range of* n_c*. (b) Zoom in to* n_c *between* 25 *and* 120*. (c) Zoom in further to* n_c *between* 60 *and* 100*.*

on this region. We see in panel (c) that the Hill estimator is close to 2.2 when $n(c)$ is between 60 and 100, and we will take 2.2 as the Hill estimate. Thus, the Hill estimate is similar to the regression estimate (1.975) of the tail index.

The advantage of the regression estimate is that one can use the linearity of the plots of $\{(\log(k/n), -R_{(k)})\}_{k=1}^{m}$ for different m to guide the choice of m, which is analogous to $n(c)$. A linear plot indicates a polynomial tail. In contrast, the Hill plot checks for the stability of the estimator and does not give a direct assessment whether or not the tail is polynomial.

□

19.7 Pareto Distributions

The Pareto distribution with location parameter $c > 0$ and shape parameter $a > 0$ has density

$$f(y; a, c) = \begin{cases} ac^a \, y^{-(a+1)}, & y > c, \\ 0, & \text{otherwise.} \end{cases} \tag{19.35}$$

The expectation is $ac/(a-1)$ if $a > 1$ and $+\infty$ otherwise. The Pareto distribution has a polynomial tail and, in fact, a polynomial tail is often called a Pareto tail.

Equation (19.30) states that the loss, conditional on being above $|c|$, has a Pareto distribution. A property of the Pareto distribution that was exploited before [see (19.23)] is that if Y has a Pareto distribution with parameters a and c and if $d > c$, then the conditional distribution of Y, given that $Y > d$, is Pareto with parameters a and d.

19.8 Choosing the Horizon and Confidence Level

The choice of horizon and confidence coefficient are somewhat interdependent and depend on the eventual use of the VaR estimate. For shorter horizons such as one day, a large α (small confidence coefficient $= 1 - \alpha$) would result in frequent losses exceeding VaR. For example, $\alpha = 0.05$ would result in a loss exceeding VaR approximately once per month since there are slightly more than 20 trading days in a month. Therefore, we might wish to uses smaller values of α with a shorter horizon.

One should be wary, however, of using extremely small values of α, such as, values less than 0.01. When α is very small, then VaR and, especially, ES are impossible to estimate accurately and are very sensitive to assumptions about the left tail of the return distribution. As we have seen, it is useful to create bootstrap confidence intervals to indicate the amount of precision in the VaR and ES estimates. It is also important to compare estimates based on different tail assumptions as in Figure 19.5, for example, where the three estimates of VaR are increasingly dissimilar as α decreases below 0.01.

There is, of course, no need to restrict attention to only one horizon or confidence coefficient. When VaR is estimated parametrically and i.i.d. normally distributed returns are assumed, then it is easy to reestimate VaR with different horizons. Suppose that $\widehat{\mu}_P^{\text{1day}}$ and $\widehat{\sigma}_P^{\text{1day}}$ are the estimated mean and standard deviation of the return for one day. Assuming only that returns are i.i.d., the mean and standard deviation for M days are

$$\widehat{\mu}_P^{\text{M days}} = M\widehat{\mu}_P^{\text{1 day}} \tag{19.36}$$

and

$$\widehat{\sigma}_P^{\text{M days}} = \sqrt{M}\widehat{\sigma}_P^{\text{1 day}}. \tag{19.37}$$

Therefore, if one assumes further that the returns are normally distributed, then the VaR for M days is

$$\text{VaR}_P^{\text{M days}} = -S \times \left\{ M\widehat{\mu}_P^{\text{1 day}} + \sqrt{M}\varPhi^{-1}(\alpha)\widehat{\sigma}_P^{\text{1 day}} \right\}, \tag{19.38}$$

where S is the size of the initial investment. The power of equation (19.38) is, for example, that it allows one to change from a daily to a weekly horizon without reestimating the mean and standard deviation with weekly instead of daily returns. Instead, one simply uses (19.38) with $M = 5$. The danger in using (19.38) is that it assumes normally distributed returns and no autocorrelation or GARCH effects (volatility clustering) of the daily returns. If there is positive autocorrelation, then (19.38) underestimates the M-day VaR. If there are GARCH effects, then (19.38) gives VaR based on the marginal distribution, but one should be using VaR based on the conditional distribution given the current information set.

If the returns are not normally distributed, then there is no simple analog to (19.38). For example, if the daily returns are i.i.d., t-distributed then one cannot simply replace the normal quantile $\Phi^{-1}(\alpha)$ in (19.38) by a t-quantile. The problem is that the sum of i.i.d. t-distributed random variables is not itself t-distributed. Therefore, if the daily returns are t-distributed then the sum M daily returns is not t-distributed. However, for large values of M and i.i.d. returns, the sum of M independent returns will be close to normally distributed by the central limit theorem, so (19.38) could be used for large M even if the returns are not normally distributed.

19.9 VaR and Diversification

A serious problem with VaR is that it may *discourage* diversification. This problem was studied by Artzner, Delbaen, Eber, and Heath (1997, 1999), who ask the question, what properties can reasonably be required of a risk measure? They list four properties that any risk measure should have, and they call a risk measure *coherent* if it has all of them.

One property among the four that is very desirable is *subadditivity*. Let $\mathfrak{R}(P)$ be a risk measure of a portfolio P, for example, VaR or ES. Then $\mathfrak{R}$ is said to be subadditive, if for any two portfolios P_1 and P_2, $\mathfrak{R}(P_1 + P_2) \leq \mathfrak{R}(P_1) + \mathfrak{R}(P_2)$. Subadditivity says that the risk for the combination of two portfolios is at most the sum of their individual risks, which implies that diversification reduces risk or at least does not increase risk. For example, if a bank has two traders, then the risk of them combined is less than or equal to the sum of their individual risks if a subadditive risk measure is used. Subadditivity extends to more than two portfolios, so if $\mathfrak{R}$ is subadditive, then for m portfolios, $P_1, \ldots, P_m$,

$$\mathfrak{R}(P_1 + \cdots + P_m) \leq \mathfrak{R}(P_1) + \cdots + \mathfrak{R}(P_m).$$

Suppose a firm has 100 traders and monitors the risk of each trader's portfolio. If the firm uses a subadditive risk measure, then it can be sure that the total risk of the 100 traders is at most the sum of the 100 individual risks. Whenever this sum is acceptable, there is no need to compute the risk measure for the entire firm. If the risk measure used by the firm is not subadditive, then there is no such guarantee.

Unfortunately, as the following example shows, VaR is *not* subadditive and therefore is incoherent. ES is subadditive, which is a strong reason for preferring ES to VaR.

Example 19.8. An example where VaR is not subadditive

This simple example has been designed to illustrate that VaR is not subadditive and can discourage diversification. A company is selling par \$1000 bonds with a maturity of one year that pay a simple interest of 5% so that the bond pays \$50 at the end of one year if the company does not default. If the bank defaults, then the entire \$1000 is lost. The probability of no default is 0.96. To make the loss distribution continuous, we will assume that the loss is $N(-50, 1)$ with probability 0.96 and $N(1000, 1)$ with probability 0.04. The main purpose of making the loss distribution continuous is to simplify calculations. However, the loss would be continuous, for example, if the portfolio contained both the bond and some stocks. Suppose that there is a second company selling bonds with exactly the same loss distribution and that the two companies are independent.

Consider two portfolios. Portfolio 1 buys two bonds from the first company and portfolio 2 buys one bond from each of the two companies. Both portfolios have the same expected loss, but the second is more diversified. Let $\Phi(x; \mu, \sigma^2)$ be the normal CDF with mean μ and variance σ^2. For portfolio 1, the loss CDF is

$$0.96\,\Phi(x; 2000, 4) + 0.04\,\Phi(x; -100, 4),$$

while for portfolio 2, by independence of the two companies, the loss distribution CDF is

$$0.96^2\,\Phi(x; 2000, 2) + 2(0.96)(0.04)\,\Phi(x; 950, 2) + 0.04^2\,\Phi(x; -100, 2).$$

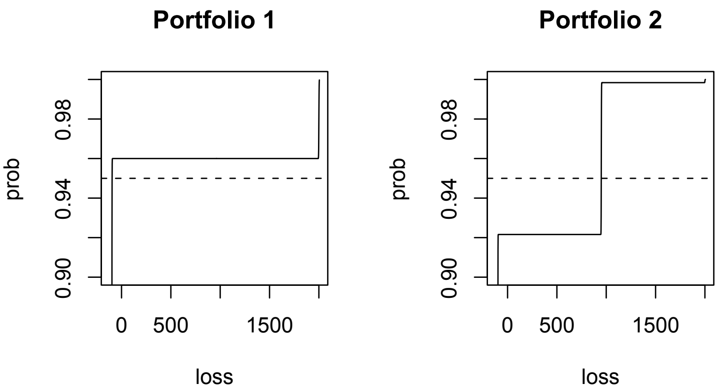

Fig. 19.7. *Example where VaR discourages diversification. Plots of the CDF of the loss distribution. VaR(*0.05*) is the loss at which the CDF crosses the horizontal dashed line at* 0.95.

We should expect the second portfolio to seem less risky, but VaR(0.05) indicates the opposite. Specifically, VaR(0.05) is -95.38 and 949.53 for portfolios 1 and 2, respectively. Notice that a negative VaR means a negative loss (positive revenue). Therefore, portfolio 1 is much less risky than portfolio 2, at least as measured by VaR(0.05). For each portfolio, VaR(0.05) is shown in Figure 19.7 as the loss at which the CDF crosses the horizontal dashed line at 0.95.

Notice as well that which portfolio has the highest value of VaR(α) depends heavily on the values of α. When α is below the default probability, 0.04, portfolio 1 is more risky than portfolio 2.

□

Although VaR is often considered the industry standard for risk management, Artzner, Delbaen, Eber, and Heath (1997) make an interesting observation. They note that when setting margin requirements, an exchange should use a subadditive risk measure so that the aggregate risk due to all customers is guaranteed to be smaller than the sum of the individual risks. Apparently, no organized exchanges use quantiles of loss distributions to set margin requirements. Thus, exchanges may be aware of the shortcomings of VaR, and VaR is not the standard for measuring risk within exchanges.

19.10 Bibliographic Notes

Risk management is an enormous subject and we have only touched upon a few aspects, focusing on statistical methods for estimating risk. We have not considered portfolios with bonds, foreign exchange positions, interest rate derivatives, or credit derivatives. We also have not considered risks other than market risk or how VaR and ES can be used for risk management. To cover risk management thoroughly requires at least a book-length treatment of that subject. Fortunately, excellent books exist, for example, Dowd (1998), Crouhy, Galai, and Mark (2001), Jorion (2001), and McNeil, Frey, and Embrechts (2005). The last has a strong emphasis on statistical techniques, and is recommended for further reading along the lines of this chapter. Generalized Pareto distributions were not covered here but are discussed in McNeil, Frey, and Embrechts.

Alexander (2001), Hull (2003), and Gourieroux and Jasiak (2001) have chapters on VaR and risk management. The semiparametric method of estimation based on the assumption of a polynomial tail and equation (19.20) are from Gourieroux and Jasiak (2001). Drees, de Haan, and Resnick (2000) and Resnick (2001) are good introductions to Hill plots.

19.11 References

Alexander, C. (2001) *Market Models: A Guide to Financial Data Analysis*, Wiley, Chichester.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1997) Thinking coherently. *RISK*, **10**, 68–71.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999) Coherent measures of risk. *Mathematical Finance*, **9**, 203–238.

Crouhy, M., Galai, D., and Mark, R. (2001) *Risk Management*, McGraw-Hill, New York.

Drees, H., de Haan, L., and Resnick, S. (2000) How to make a Hill plot, *Annals of Statistics*, **28**, 254–274.

Dowd, K. (1998) *Beyond Value At Risk*, Wiley, Chichester.

Gourieroux, C., and Jasiak, J. (2001) *Financial Econometrics*, Princeton University Press, Princeton, NJ.

Hull, J. C. (2003) *Options, Futures, and Other Derivatives*, 5th ed., Prentice-Hall, Upper Saddle River, NJ.

Jorion, P. (2001) *Value At Risk*, McGraw-Hill, New York.

McNeil, A. J., Frey, R., and Embrechts, P. (2005) *Quantitative Risk Management*, Princeton University Press, Princeton, NJ.

Resnick, S. I. (2001) *Modeling Data Networks*, School of Operations Research and Industrial Engineering, Cornell University, Technical Report #1345.

19.12 R Lab

19.12.1 VaR Using a Multivariate-*t* Model

Run the following code to create a data set of returns on two stocks, DATGEN and DEC.

```
library("fEcofin")
library(mnormt)
Berndt = berndtInvest[,5:6]
names(Berndt)
```

Problem 1 *Fit a multivariate-t model to* `Berndt`*; see Section 7.14.3 for an example of fitting such a model. What are the estimates of the mean vector, DF, and scale matrix? Include your* R *program with your work. Include your* R *code and output with your work.*

Problem 2

(a) What is the distribution of the return on a $100,000 portfolio that is 30% invested in DATGEN and 70% invested in DEC? Include your R *code and output with your work.*

(b) Find $\widehat{VaR}^t(0.05)$ *and* $\widehat{ES}^t(0.05)$ *for this portfolio.*

Problem 3 *Use the model-free bootstrap to find a basic percentile bootstrap confidence interval for VaR(0.05) for this portfolio. Use a 90% confidence coefficient for the confidence interval. Use 250 bootstrap resamples. This amount* *of resampling is not enough for a highly accurate confidence interval, but will give a reasonably good indication of the uncertainty in the estimate of VaR(0.05), which is all that is really needed.*

Also, plot kernel density estimates of the bootstrap distribution of DF and $VaR^t(0.05)$. Do the densities appear Gaussian or skewed? Use a normality test to check if they are Gaussian.

Include your `R` *code, plots, and output with your work.*

Problem 4 *This problem uses the variable DEC. Estimate the left tail index using the Hill estimator. Use a Hill plot to select n_c. What is your choice of n_c? Include your* `R` *code and plot with your work.*

19.13 Exercies

1. This exercise uses daily BMW returns in the `bmwRet` data set in the `fEcofin` package. Assume that the returns are i.i.d., even though there may be some autocorrelation and volatility clustering is likely.
   (a) Compute nonparametric estimates of VaR(0.01, 24 hours) and ES(0.01, 24 hours).
   (b) Compute parametric estimates of VaR(0.01, 24 hours) and ES(0.01, 24 hours) assuming that the returns are normally distributed.
   (c) Compute parametric estimates of VaR(0.01, 24 hours) and ES(0.01, 24 hours) assuming that the returns are t-distributed.
   (d) Compare the estimates in (a), (b), and (c). Which do you feel are most realistic?
2. Assume that the loss distribution has a polynomial tail and an estimate of a is 3.1. If VaR(0.05) = \$252, what is VaR(0.005)?
3. Find a source of stock price data on the Internet and obtain daily prices for a stock of your choice over the last 1000 days.
   (a) Assuming that the loss distribution is t, find the parametric estimate of VaR(0.025, 24 hours).
   (b) Find the nonparametric estimate of VaR(0.025, 24 hours).
   (c) Use a t-plot to decide if the normality assumption is reasonable.
   (d) Estimate the tail index assuming a polynomial tail and then use the estimate of VaR(0.025, 24 hours) from part (a) to estimate VaR(0.0025, 24 hours).
4. This exercise uses daily data in the `msft.dat` data set in the `fEcofin` package. Use the closing prices to compute daily returns. Assume that the returns are i.i.d., even though there may be some autocorrelation and volatility clustering is likely. Use the model-free bootstrap to find 95% confidence intervals for parametric estimates of VaR(0.005, 24 hours) and ES(0.005, 24 hours) assuming that the returns are t-distributed.

5. Suppose the risk measure $\mathfrak{R}$ is VaR(α) for some α. Let P_1 and P_2 be two portfolios whose returns have a joint normal distribution with means μ_1 and μ_2, standard deviations σ_1 and σ_2, and correlation ρ. Suppose the initial investments are S_1 and S_2. Show that $\mathfrak{R}(P_1+P_2) \leq \mathfrak{R}(P_1)+\mathfrak{R}(P_2)$.[1]
6. The problem uses daily stock price data in the file `Stock_FX_Bond.csv` on the book's website. In this exercise, use only the first 500 prices on each stock. The following `R` code reads the data and extracts the first 500 prices for five stocks. "`AC`" in the variables' names means "adjusted closing" price.

```
dat = read.csv("Stock_FX_Bond.csv",header=T)
prices = as.matrix(dat[1:500,c(3,5,7,9,11)])
```

   (a) What are the sample mean vector and sample covariance matrix of the 499 returns on these stocks?
   (b) How many shares of each stock should one buy to invest $50 million in an equally weighted portfolio? Use the prices at the end of the series, e.g., `prices[,500]`.
   (c) What is the one-day VaR(0.1) for this equally weighted portfolio? Use a parametric VaR assuming normality.
   (d) What is the five-day Var(0.1) for this portfolio? Use a parametric VaR assuming normality. You can assume that the daily returns are uncorrelated.

[1] This result shows that VaR is subadditive on a set of portfolios whose returns have a joint normal distribution, as might be true for portfolios containing only stocks. However, portfolios containing derivatives or bonds with nonzero probabilities of default generally do not have normally distributed returns.