Lecture 7
Risk Management

CHEN Ying
FE5209 Financial Econometrics

Outline

- Risk measures: Value-at-Risk (VaR) and Expected Shortfall (ES)
- Estimate VaR and ES.
- Tail index under polynomial tails assumption
- VaR and diversification

Readings
SDA chapter 19

Risks

There are many different types of risk:

- Market Risk - risk associated with fluctuations in value of traded assets *(e.g. financial crisis)*
- Credit Risk - risk associated with uncertainty that debtors will honour their financial obligations
- Operational Risk - human error, IT failure, dishonesty, natural disaster etc. *(e.g. fat-finger in 2006 costed 40 billion Yen of a Japanese firm. A junior employee at Deutsche Bank caused by mistake a payment to a US hedge fund of $6bn in 2015)*
- Liquidity risk - potential extra cost of liquidating a position because buyers are difficult to locate *(e.g. Amaranth Advisors lost roughly $6bn in the natural gas futures market back in September 2006, due to liquidity. Long-Term Capital Management was bailed out in 1998 due to funding and asset liquidity.)*

Risks are represented by random variables mapping unforeseen future states of the world into values representing profits and losses. The risks which interest us are aggregate risks. In general we consider a portfolio which might be

- a collection of stocks and bonds;
- a book of derivatives;
- a collection of risky loans;
- a financial institution's overall position in risky assets.

Losses

Consider a portfolio and let V_t denote its value at time t; we assume this random variable is observable at time t.

Suppose we look at risk from perspective of time t and we consider the time period $[t, t + 1]$. The value V_{t+1} at the end of the time period is unknown to us.
The distribution of $(V_{t+1} - V_t)$ is known as the profit-and-loss or P&L distribution.

We denote the loss by $L_{t+1} = -(V_{t+1} - V_t)$.
By this convention, losses will be positive numbers and profits negative.
We refer to the distribution of L_{t+1} as the loss distribution.

Risk measures

Risk measures attempt to quantify the riskiness of a portfolio over the holding period T. Denote the distribution function of the loss L by F_L so that

$$P(L \leq x) = F_L(x).$$

Primary risk measure: Value at Risk defined as

$$VaR_\alpha = q_\alpha(F_L) \quad or \quad P(L \geq VaR_\alpha) = \alpha$$

VaR describe the right tail of the loss distribution of L_{t+1} (or the left tail of the P&L). It is the αth upper quantile of L. VaR is a bound such that the loss over the horizon is less than this bound with probability equal to the confidence coefficient.

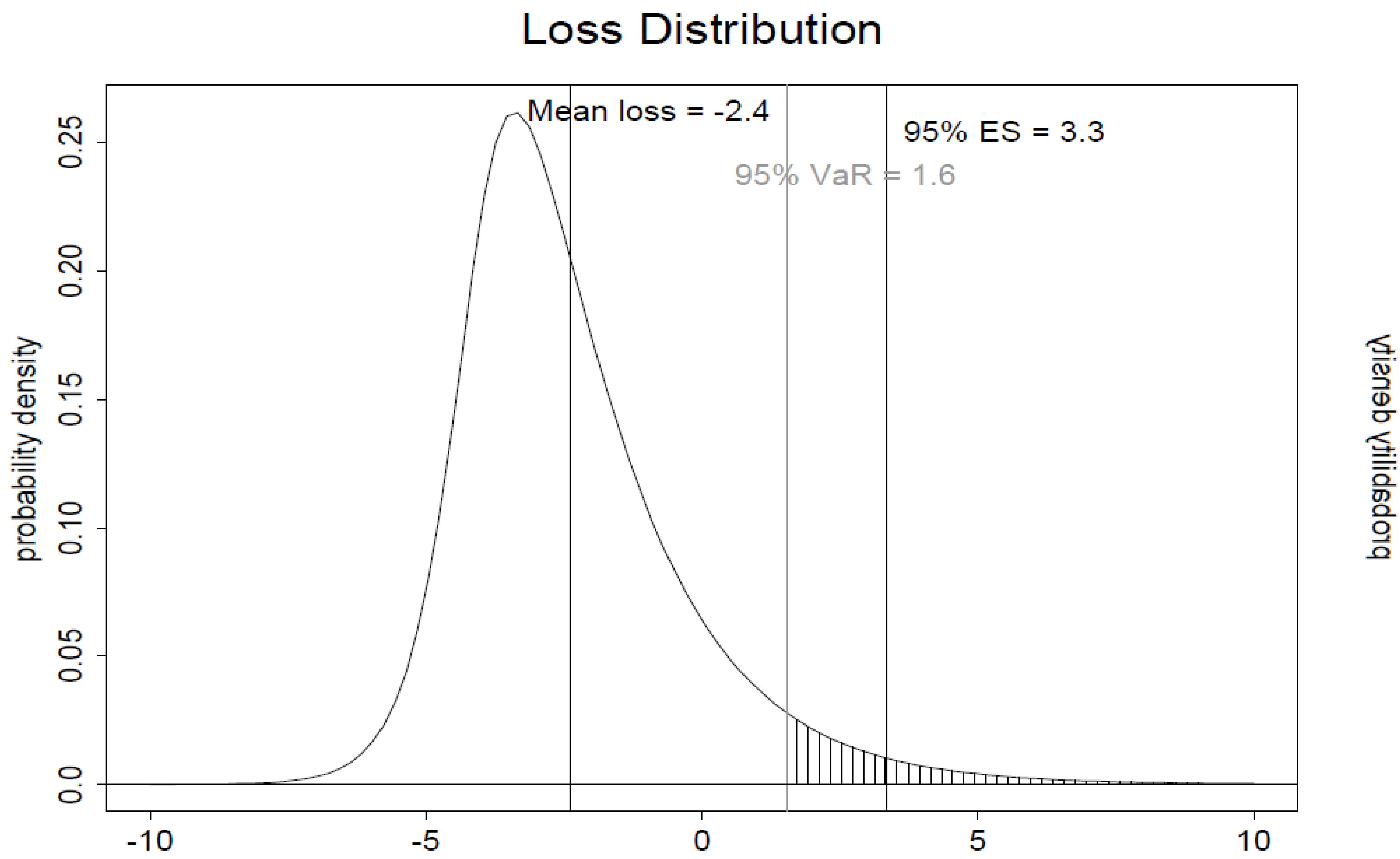
For example, if the horizon is one week, the confidence coefficient is 99% (so $\alpha = 0.01$), and the VaR is $5 million, then there is only a 1% chance of a loss exceeding $5 million over the next week.

Alternative risk measure: Expected shortfall defined as

$$ES_\alpha = E(L|L \geq VaR_\alpha)$$

i.e. the average loss when VaR is exceeded. ES gives information about frequency and size of large losses. ES is called by a variety of names: expected shortfall, the expected loss given a tail event, tail loss, and shortfall.

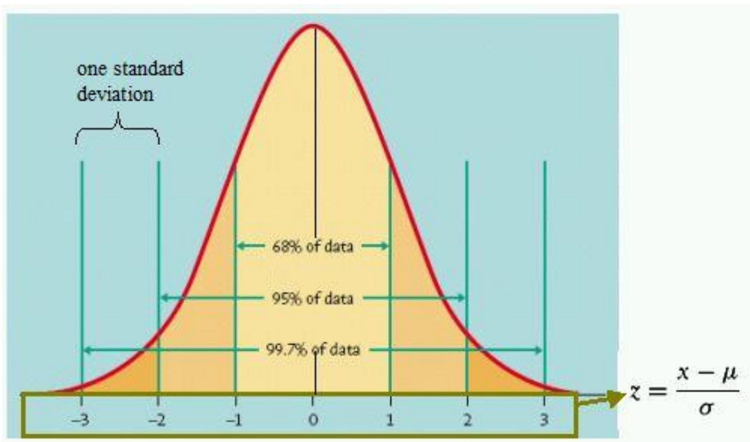
Risk measures



Example: 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?

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.



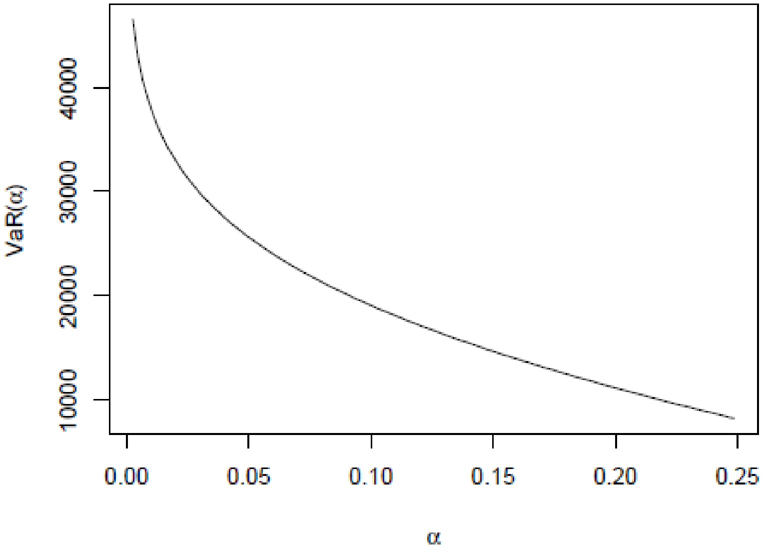


Fig. 19.1. VaR(α) *for* $0.025 < \alpha < 0.25$ *when the loss distribution is normally distributed with mean* -4000 *and standard deviation* $18{,}000$.

Variance-Covariance method

We assume risk factors has a multivariate normal distribution. We assume that the linearized loss in terms of risk factors is a sufficiently accurate approximation of the loss.

Assume risk factor $X_t \sim N(\mu, \sigma^2)$. The loss distribution is approximated by the distribution of $N(-\mu, \sigma^2)$.

- The mean vector and covariance matrix are estimated from data $X_1, \dots, X_T$ to give estimates $\hat{\mu}$ and $\hat{\sigma}$.
- Inference about the loss distribution is made using distribution $N(-\hat{\mu}, \hat{\sigma}^2)$.

Estimates of the risk measures VaR and ES are calculated from the estimated loss distribution.

VaR is estimated by $\widehat{VaR}_\alpha = S \times \{-\hat{\mu} + \hat{\sigma}\Phi^{-1}(1-\alpha)\}$

ES is estimated by $\widehat{ES}_\alpha = S \times \{-\hat{\mu} + \hat{\sigma}\frac{\phi\left(\Phi^{-1}(1-\alpha)\right)}{\alpha}\}$.

where ϕ is standard normal density, Φ is the cdf and $z_\alpha = \Phi^{-1}(1-\alpha)$ is the α-upper quantile of the standard normal distribution.

$$ES = E(L_{t+1}|L_{t+1} > VaR_\alpha) = E(L_{t+1}|L_{t+1} > -\mu_t + \sigma_t\ z_\alpha)$$

$$= -\mu_t + \sigma_t\, E\left(\frac{L_{t+1}-(-\mu_t)}{\sigma_t}\middle| \frac{L_{t+1}-(-\mu_t)}{\sigma_t} > z_\alpha\right)$$

Note $\frac{L_{t+1}-(-\mu_t)}{\sigma_t}$ is standard normal. It's conditional expectation is

$$\mathrm{E}(Z_{t+1}|Z_{t+1} > u) = \frac{\phi(u)}{1-\Phi(u)}.$$

Extensions Variance-Covariance with t-distribution

Instead of assuming normal risk factors, the method could be easily adapted to use multivariate Student t risk factors or multivariate hyperbolic risk factors, without sacrificing tractability.

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 VaR is $\widehat{VaR}_\alpha = S \times \{-\hat{\mu} + \hat{\lambda} q_{\alpha,t}(\hat{\nu})\}$

The ES is $\widehat{ES}_\alpha = S \times \left\{-\hat{\mu} + \hat{\lambda}\left(\frac{f_{\hat{\nu}}\{F_{\hat{\nu}}^{-1}(\alpha)\}}{\alpha}\left[\frac{\hat{\nu}+\{F_{\hat{\nu}}^{-1}(\alpha)\}^2}{\hat{\nu}-1}\right]\right)\right\}$

$$f_\nu = p(x|\nu,\mu,\lambda) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}\left(\frac{\lambda}{\pi\nu}\right)^{\frac{1}{2}}\left(1 + \frac{\lambda(x-\mu)^2}{2}\right)^{-\frac{\nu+1}{2}}$$

$$E(x) = \mu$$

$$\sigma = \lambda\sqrt{\nu/(\nu-2)}$$

Example: VaR and ES for a position in an S&P 500 index fund

Suppose that you hold a $20,000 position in an S&P 500 index fund, and that you want a 24-hour VaR at 95% confidence.

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. Black Monday, with a log return of -.23, occurs near the beginning of the shortened time series used in this example.

We will assume that the returns are i.i.d. with a *t*-distribution.

The *t*-distribution was fit using R's fitdistr function and the estimates were $\hat{\mu} = 0.000689$, $\hat{\lambda} = 0.007164$, and $\hat{\nu} = 2.984$. The estimated standard deviation is $\hat{\sigma} = \hat{\lambda}\sqrt{\hat{\nu}/(\hat{\nu}-2)} = 0.01248$. The 0.05-quantile of the *t*-distribution with 2.984 degrees of freedom is -2.3586.

Therefore, $\widehat{VaR}_{5\%} = -20000 \times \{0.000689 - (2.3586)(0.007164)\} = \$323.42$.

The parametric estimate of $\widehat{ES}_{\alpha} = \$543.81$.

- 7_VaR_SP500.R

Historical simulation method

Non-parametric: 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. The first step is to sort the risk factors e.g. returns such that

$$x_{(1)} < x_{(2)} < \ldots < x_{(T)}$$

The VaR is estimated using $x_{\lfloor \alpha T \rfloor}$ or alternatively $x_{\lceil \alpha T \rceil}$ or an average of the two In other words, the estimate of the VaR is the α -quantile of the return distribution, which is the α -upper quantile of the loss distribution.

If S is the size of the current position, then the nonparametric estimate of VaR is

$$\widehat{VaR}_\alpha = -S \times \hat{q}_\alpha$$

where $\hat{q}_\alpha$ is the empirical quantile of the historical returns. The minus sign converts revenue (return times initial investment) to a loss.

To estimate ES, let $R_1, \ldots, R_T$ be the historic returns and define $L_t = -S \times R_t$. Then

$$\widehat{ES}_\alpha = E(L|L \geq VaR_\alpha) = \frac{\sum_{t=1}^{T} L_t I(L_t > \widehat{VaR}_\alpha)}{\sum_{t=1}^{T} I(L_t > \widehat{VaR}_\alpha)} = -S \times \frac{\sum_{t=1}^{T} R_t I(R_t < \hat{q}_\alpha)}{\sum_{t=1}^{T} I(R_t < \hat{q}_\alpha)}$$

which is the average of all L_t exceeding $\widehat{VaR}_\alpha$. Here $I\,(.)$is the indicator.

Example: VaR and ES for a position in an S&P 500 index fund

This example uses the S&P500 data set as in the previous example so that parametric and nonparametric estimates can be compared.

Suppose that you hold a $20,000 position in an S&P 500 index fund, and that you want a 24-hour VaR at 95% confidence. The 0.05 quantile of the returns computed by R's quantile function is -.0169. In other words, a daily return of -.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 -.0169 on a $20,000 investment yields a revenue of -$337.43, and therefore the estimated $\widehat{VaR}_{5\%}$(24 hours) is $337.43.
$\widehat{ES}_{5\%}$ is obtained by averaging all returns below -.0169 and multiplying this average by -20,000. The result is $\widehat{ES}_{5\%}$ = $619.3.

- 7_VaR_SP500.R

Example: VaR and ES for a position in an S&P 500 index fund

The t-distributed variance covariance estimate ($\widehat{VaR}_{5\%}$ is \$323.42 and $\widehat{ES}_{5\%}$ = \$ 543.81) are smaller than the nonparametric estimates ($\widehat{VaR}_{5\%}$ is \$337.43 and $\widehat{ES}_{5\%}$ = \$619.3).

The reason the two sets of 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.

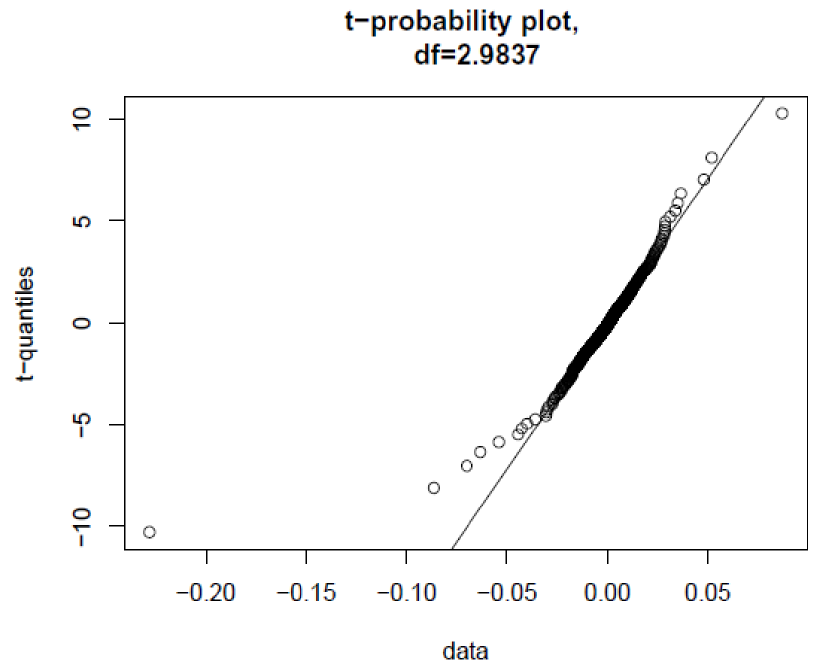


Fig. 19.2. *t-plot of the S&P 500 returns used in Examples 19.2 and 19.3. The deviations from linearity in the tails, especially the left tail, indicate that the t-distribution does not fit the data in the extreme tails. The reference line goes through the first and third quartiles. The t-quantiles use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

Estimating VaR and ES using ARMA GARCH models

Equity returns typically have a small amount of autocorrelation and a greater amount of volatility clustering. We use ARMA/GARCH models so that VaR and ES can adjust to periods of high or low volatility.

Assume that we have *T* returns, $X_1, \ldots, X_T$ and we estimate VaR and ES for the next return X_{T+1}. Let $\hat{\mu}_{T+1|T}$ and $\hat{\sigma}_{T+1|T}$ be the estimated conditional mean and variance of tomorrow's return conditional on the current information set.

We also assume that X_{T+1} has a conditional *t*-distribution with ν degrees of freedom. After fitting an ARMA/GARCH model, we have estimates of $\hat{\nu}$,$\hat{\mu}_{T+1|T}$ and $\hat{\sigma}_{T+1|T}$. The estimated conditional scale parameter is $\hat{\lambda}_{T+1|T} = \sqrt{\frac{\hat{\nu}-2}{\hat{\nu}}}\,\hat{\sigma}_{T+1|T}$

The VaR is $\widehat{VaR}_\alpha = S \times \{-\hat{\mu}_{T+1|T} + \hat{\lambda}_{T+1|T} q_{\alpha,t}(\hat{\nu})\}$

The ES is $\widehat{ES}_\alpha = S \times \left\{-\hat{\mu}_{T+1|T} + \hat{\lambda}_{T+1|T}\left(\frac{f_{\hat{\nu}}\{F_{\hat{\nu}}^{-1}(\alpha)\}}{\alpha}\left[\frac{\hat{\nu}+\{F_{\hat{\nu}}^{-1}(\alpha)\}^2}{\hat{\nu}-1}\right]\right)\right\}$

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

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 significant, so a GARCH(1,1) was used for estimation of VaR and ES.

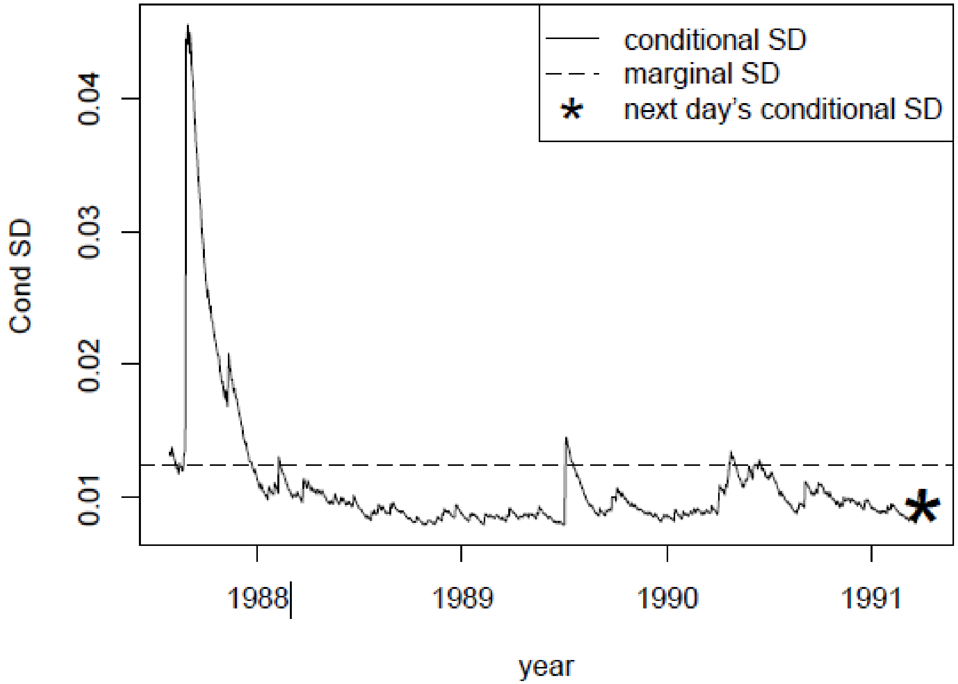
	Estimate	Std. Error	t value	Pr(>|t|)	
mu	7.147e-04	2.643e-04	2.704	0.00685	**
omega	2.833e-06	9.820e-07	2.885	0.00392	**
alpha1	3.287e-02	1.164e-02	2.824	0.00474	**
beta1	9.384e-01	1.628e-02	57.633	< 2e-16	***
shape	4.406e+00	6.072e-01	7.256	4e-13	***

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 ($323.42 and $543.81), because the conditional standard deviation used here (0.00950) is smaller than the marginal standard deviation ($\hat{\sigma} = \hat{\lambda}\sqrt{\hat{\nu}/(\hat{\nu}-2)} = 0.01248$) used.

7_VaR_SP500_GARCH.R

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



The VaR and ES estimates using the GARCH model are considerably smaller than the parametric estimates ($323.42 and $543.81), because the conditional standard deviation used here (0.00950) is smaller than the marginal standard deviation (0.01248) used.

Statistical methods for market risk

- Nonparametric: Historical Simulation Method
- Parametric: Variance-Covariance Method
- Semiparametric: Monte Carlo Simulation Method

Nonparametric vs. Parametric

Nonparametric methods:

+ Easy to implement. No statistical estimation of the loss distribution necessary.

- It may be difficult to collect sufficient quantities of relevant, synchronized data for all risk factors. Historical data may not contain examples of extreme scenarios.

Parametric methods:

+ In contrast to the nonparametric methods, parametric estimation may allow the use of GARCH models to adapt the risk measures to the current estimate of volatility. It offers analytical solution with no simulation.

- Linearization may be crude approximation. Assumption of distribution e.g. normality may seriously underestimate tail of loss distribution.

Monte Carlo simulation method

We estimate the loss distribution under some explicit parametric model for risk factor X_{T+1}. We make inference about L using Monte Carlo methods, which involves simulation of new risk factor data.

1. With the help of the historical risk factor data $X_1, \ldots, X_T$ calibrate a suitable statistical model for risk factor changes and simulate m new data $\tilde{X}_{T+1}^{(1)}, \ldots, \tilde{X}_{T+1}^{(m)}$ from this model.
2. Construct the Monte Carlo data's losses
3. Make inference about loss distribution and risk measures using the simulated data $\tilde{L}_1, \ldots, \tilde{L}_m$. We have similar possibilities as for historical simulation.

+ Very general. No restriction in our choice of distribution for X_t
- Can be very time consuming if loss operator is difficult to evaluate, which depends on size m and complexity of portfolio.

Note that MC approach does not address the problem of determining the distribution of risk factor X_{T+1}.

VaR and ES for a portfolio

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.

With a portfolio of stocks, means, variances, and covariances of returns could be estimated directly from a sample of returns.

Assume risk factor $X_t \sim N_d(\mu, \Sigma) \Rightarrow w' X_t \sim N(w'\mu, w'\Sigma \mathrm{w})$. The loss distribution is approximated by the distribution of $N(-w'\mu, w'\Sigma\mathrm{w})$.

- ❑ The mean vector and covariance matrix are estimated from data $X_1, \ldots, X_T$ to give estimates $\hat{\mu}$ and $\hat{\Sigma}$.
- ❑ Inference about the loss distribution is made using distribution $N\left(-w'\hat{\mu}, w'\hat{\Sigma}\mathrm{w}\right)$.

Estimates of the risk measures VaR and ES are calculated from the estimated loss distribution.

VaR is estimated by $\widehat{VaR}_\alpha = S \times \{-w'\hat{\mu} + \sqrt{w'\hat{\Sigma}\mathrm{w}}\Phi^{-1}(1-\alpha)\}$.

ES is estimated by $\widehat{ES}_\alpha = S \times \left\{-w'\hat{\mu} + \sqrt{w'\hat{\Sigma}\mathrm{w}}\frac{\phi\left(\Phi^{-1}(1-\alpha)\right)}{\alpha}\right\}$.

Example: VaR and ES for the three stocks in the CRSP

Consider the data set CRSP with four variables -- returns on GE, IBM, Mobil, and the CRSP index. The 3 stock returns can be modeled as a multivariate *t*-distribution with 5.81 degrees of freedom.

The estimated mean was $\hat{\mu} = (\ 0.0008584\ 0.0003249\ 0.0006162\ )^T$
and the estimated covariance matrix was

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

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

$$\hat{\mu}_P = \hat{\mu}^T w = 0.0005998$$

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

$$\hat{\Sigma}_P = w^T \hat{\Sigma} w = 0.008455.$$

Example: VaR and ES for the three stocks in the CRSP

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 $\hat{\nu} = 5.81$, is

$\hat{\lambda}_P = \sqrt{\frac{\hat{\nu}-2}{\hat{\nu}}} \times 0.008455 = 0.006847.$

Therefore,

$$\widehat{VaR}_{5\%} = S \times \left\{-\hat{\mu}_P + \hat{\lambda}_P q_{5\%,t}(\hat{\nu})\right\} = S \times 0.01278$$

$$\widehat{ES}_{5\%} = S \times \left\{-\hat{\mu}_P + \hat{\lambda}_P \left(\frac{f_{\hat{\nu}} q_{5\%,t}(\hat{\nu})}{5\%} \left[\frac{\hat{\nu} + \left\{q_{5\%,t}(\hat{\nu})\right\}^2}{\hat{\nu} - 1}\right]\right)\right\} = S \times 0.01815$$

With *S* = \$20,000, we have $\widehat{VaR}_{5\%} = \$256$ and $\widehat{ES}_{5\%} = \$363$.

Confidence interval using the Bootstrap

Suppose we have a large number, *B*, of resamples of the returns data. Then a VaR_α or ES_α estimate is computed from each resample and for the original sample. The confidence interval can be based upon either a parametric or nonparametric estimator.

Suppose that we want the confidence coefficient of the interval to be $\mathbf{1-\gamma}$. *The interval's confidence coefficient should not be confused with the confidence coefficient of VaR,* $1-\alpha$*.*

The $\gamma/2$-lower and -upper quantiles of the bootstrap estimates of $1-\gamma$ are the limits of the basic percentile method confidence intervals.

Example: Bootstrap confidence intervals

We continue the previous SP500 examples and find a confidence interval for VaR_{α} and ES_{α}. We use $\alpha = 0.05, \gamma = 0.1$, B = 5,000 resamples were taken.

The basic 90% $= 1 - \gamma$ confidence intervals for $VaR_{5\%}$ were (297, 352) and (301, 346) using nonparametric and parametric estimators respectively.
For $ES_{5\%}$ the corresponding 90% percentile confidence intervals were (487, 803) and (433, 605).

We see that there is considerable uncertainty in the risk measures, especially for ES and especially using nonparametric estimation.

The bootstrap computation took 33.3 minutes using an R program and a 2.13 GHz PentiumTM processor running under WindowsTM.

- 7_VaR_SP500.R

Estimation of VaR with polynomial tails

Flexibility vs. Accuracy:

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.

Semiparamtric approach:

+ Parametric assumption: the return density has a polynomial left tail, or equivalently that the loss density has a polynomial right-tail. $f(x) \sim Ax^{-(\mathrm{a}+1)}$ as $\mathrm{x} \to -\infty$

+ Use a nonparametric estimate of VaR_{α_0} for a *large* value of α_0 to obtain estimates of VaR_{α_1} for *small* values of α_1. It is assumed here that VaR_{α_0} and VaR_{α_1} have the same horizon *T*.

Polynomial tails: VaR

The return density $f(x)$ is assumed to have a polynomial left tail,

$$f(x) \sim Ax^{-(a+1)} \text{ as } x \to -\infty$$

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

$$P(X \le x) \sim \int_{-\infty}^{x} f(u)du = \frac{A}{a}x^{-a}, \qquad \text{as } x \to -\infty$$

and if $x_1 > 0$ and $x_2 > 0$, then

$$\frac{P(X \le -x_1)}{P(X \le -x_2)} \approx \left(\frac{x_1}{x_2}\right)^{-a}$$

Now suppose that $x_1 = \mathrm{VaR}_\alpha$ and $x_2 = \mathrm{VaR}_{\alpha_0}$ where $0 < \alpha < \alpha_0$.
Then we have

$$\frac{\alpha}{\alpha_0} = \frac{P(X \le -\mathrm{VaR}_\alpha)}{P\left(X \le -\mathrm{VaR}_{\alpha_0}\right)} \approx \left(\frac{\mathrm{VaR}_\alpha}{\mathrm{VaR}_{\alpha_0}}\right)^{-a}$$

or:

$$\mathrm{VaR}_\alpha = \mathrm{VaR}_{\alpha_0}\left(\frac{\alpha_0}{\alpha}\right)^{\frac{1}{a}}$$

Polynomial tails: ES

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

$$\mathrm{f(x)} \sim \mathrm{A|x|^{-(a+1)}} \quad \text{as } \mathrm{x} \leq c$$

where A > 0 is a constant and a > 0 is the tail index. Then, for any $\mathrm{d} \leq c$

$$\mathrm{P(X \leq d)} \sim \int_{-\infty}^{\mathrm{d}} \mathrm{f(u)du} = \frac{\mathrm{A}}{\mathrm{a}} \mathrm{|d|^{-a}},$$

and the conditional density of X given that $\mathrm{X} \leq d$ is

$$\mathrm{f}(x|\mathrm{X} \leq \mathrm{d}) = \frac{\mathrm{A|x|^{-(a+1)}}}{\mathrm{P(X \leq d)}} = \mathrm{a|d|^a}|x|^{-(\mathrm{a}+1)},$$

It follows that for $a > 1$,

$$\mathrm{E}(x|\mathrm{X} \leq \mathrm{d}) = \mathrm{a|d|^a} \int_{-\infty}^{\mathrm{d}} |x|^{-(\mathrm{a}+1)} \mathrm{dx} = \frac{\mathrm{a}}{\mathrm{a}-1} |d|,$$

If we let $d = -VaR_\alpha$, then we see that

$$ES_\alpha = \frac{\mathrm{a}}{\mathrm{a} - 1} VaR_\alpha = \frac{1}{1 - \mathrm{a}^{-1}} VaR_\alpha \quad if\ \mathrm{a} > 1$$

It enables one to estimate ES_α using an estimate of VaR_α and an estimate of the tail index a.

Example: Estimating the tail index of the S&P 500 returns

Suppose we have invested \$20,000 in an S&P 500 index fund. We will use $\alpha_0 = 0.1$. The risk measure $VaR_{0.1}(24\,hours) = \$234$ is estimated to be -\$20,000 times -0.0117, the 0.1-quantile of the 1000 returns.

We have

$$\widehat{VaR}_\alpha\,(24\,hours) = 234\left(\frac{0.1}{\alpha}\right)^{0.506}$$

$$\widehat{ES}_\alpha = 1.975/0.975\,\widehat{VaR}_\alpha = 2.026\widehat{VaR}_\alpha$$

Regression estimator of the tail Index

According to the density with polynomial tails $\mathrm{P}(\mathrm{X} \leq \mathrm{d}) = \frac{\mathrm{A}}{\mathrm{a}}|\mathrm{d}|^{-\mathrm{a}}$

$$\log\{P(X \leq x)\} = \log(Q) - \mathrm{a}\log(|x|), \qquad Q = A/\mathrm{a}$$

If $X_{(1)}, \ldots, X_{(T)}$ are the order statistics of the returns, then the number of observed returns less than or equal to $\mathrm{X}_{(k)}$ is k.

We have

$$\log(k/T) \approx \log(Q) - \mathrm{a}\log\left(-X_{(k)}\right)$$

or, rearranging we have

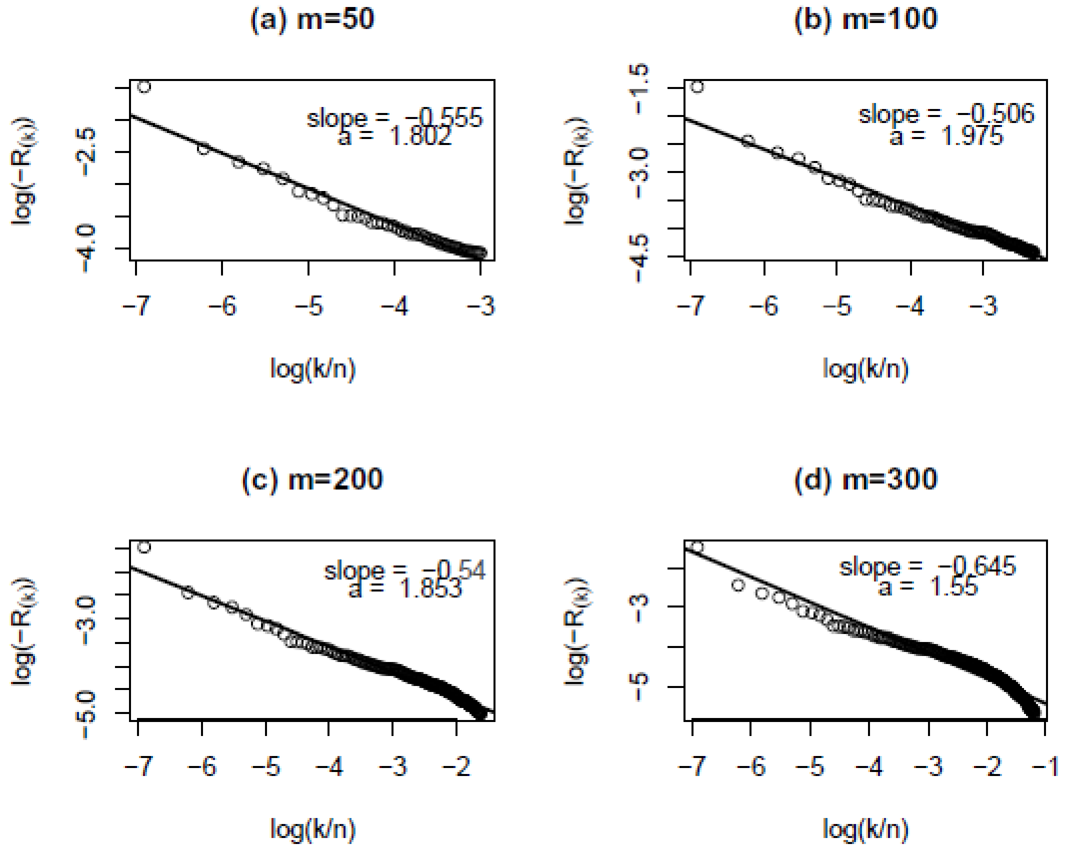
$$\log\left(-X_{(k)}\right) \approx \left(\frac{1}{\mathrm{a}}\right)\log(Q) - \left(\frac{1}{\mathrm{a}}\right)\log(k/T)$$

- The approximation is expected to be accurate only if $-X_{(k)}$ is large, which means *k* is small, perhaps only 5%, 10%, or 20% of the sample size *T*.
- It is a regression model of $\log\left(-X_{(k)}\right)$ on $\log(k/T)$, with $k = 1, \ldots, m$. The estimated slope is estimates $-1/\mathrm{a}$.

Example: Estimating the tail index of the S&P 500 returns

Consider 1000 daily S&P 500 returns.

Regression estimator of the tail index was calculated 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 slope of the line with *m* = 100 was -0.506, so $\hat{a} = 1/0.506 = 1.975$.

- 7_VaR_TailIndex.R

Hill estimator

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

$$\mathrm{f(x)} \sim \mathrm{A|x|}^{-(\mathrm{a}+1)} \text{ as } \mathrm{x} \leq c$$

Let $X_{(1)},..,X_{(T)}$ be order statistics of the returns and $T(c)$ be the number of $X_{(i)} \leq c$. The conditional density of X_i given that $X_i \leq c$ is

$$\mathrm{f}(x|\mathrm{X} \leq \mathrm{c}) = \mathrm{a|c|^a|x|}^{-(\mathrm{a}+1)}$$

Therefore, the log-likelihood for $X_{(1)},..,X_{(T(c))}$ is

$$\log\{\mathrm{L(a)}\} = \sum_{i=1}^{T(c)} \{\log(\mathrm{a}) + \mathrm{a}\log(|c|) - (\mathrm{a}+1)\log(|X_{(i)}|\}$$

FOC with respect to a gives:

$$\frac{\mathrm{T(c)}}{\mathrm{a}} = \sum_{i=1}^{T(c)} \log(|X_{(i)}/c|)$$

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

$$\hat{\mathrm{a}}^{Hill}(c) = \frac{T(c)}{\sum_{i=1}^{T(c)} \log\left(\left|\frac{X_{(i)}}{c}\right|\right)}$$

How should c be chosen?

Usually *c* is equal to one of *returns* so that $c = X_{T(c)}$, and therefore choosing *c* means choosing *T*(*c*). The choice involves a bias-variance trade off.

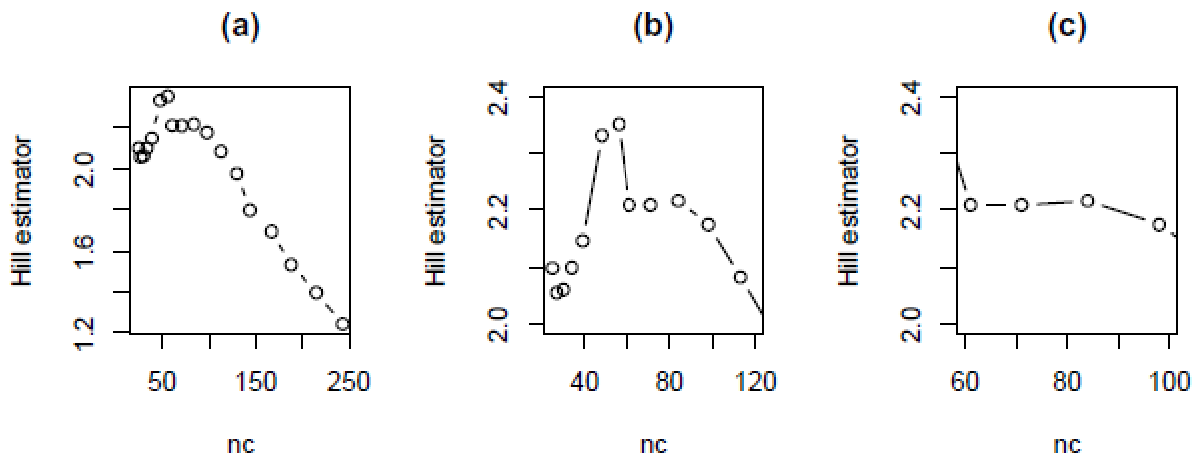
- If *T*(*c*) is too large, then $\mathrm{f(x)} \sim \mathrm{A|x|^{-(a+1)}}$ will not hold precisely for all values of $\mathrm{x} \leq c$, causing bias.
- If *T*(*c*) is too small, then there will be too few *returns* below *c* and $\hat{\mathrm{a}}^{Hill}(c)$ will be highly variable and unstable because it uses too few data.

We hope that there is a range of values of T(c) where $\hat{\mathrm{a}}^{\mathrm{Hill}}(\mathrm{c})$ is reasonably constant because it is neither too biased nor too variable.

A *Hill plot* is a plot of $\hat{\mathrm{a}}^{Hill}(c)$ versus *T*(*c*). It is used to find this range of values of *T*(*c*). In a Hill plot, where the Hill estimator is nearly constant.

Example: Estimating the tail index of the S&P 500 returns

The Hill estimator of a was also implemented. There seems to be a region of stability when *T*(*c*) is between 60 and 80. We will take 2.2 as the Hill estimate.



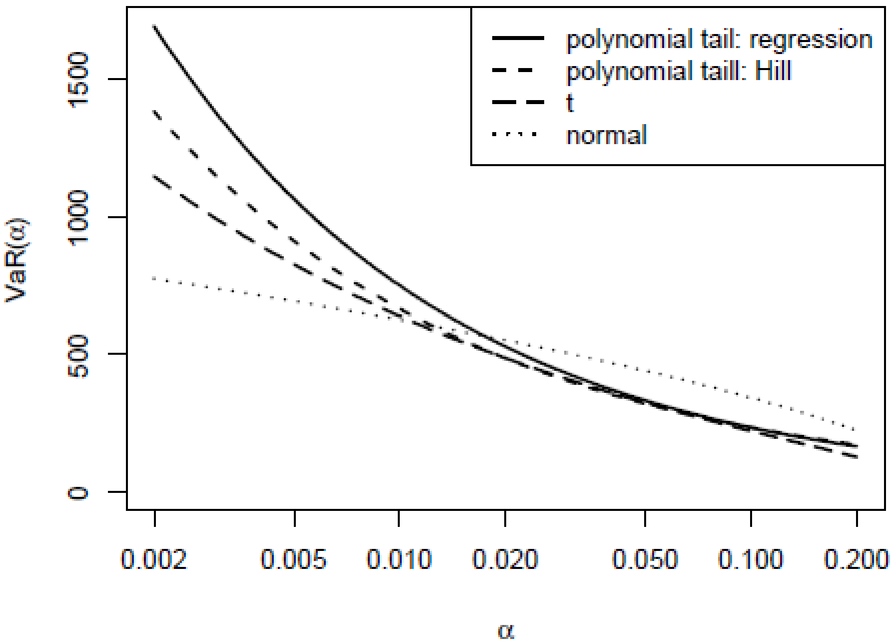
The Hill estimate is similar to the regression estimate (1.975) of the tail index.

- ❑ The advantage of the regression estimate is that one can use the linearity of the plots of $\{\log(k = T), -X(k)\}_{k=1}^{m}$ for different *m* to guide the choice of *m*, which is analogous to *T*(*c*). A linear plot indicates a polynomial tail.
- ❑ The Hill plot checks for the stability of the estimator and does not give a direct assessment whether or not the tail is polynomial.
- ❑ 7_VaR_Hill.R

Example: Estimating the tail index of the S&P 500 returns

- ❑ The return distribution has much heavier tails than a normal distribution.
- ❑ *The polynomial tail with tail index 1.975 or Hill estimator 2.2 is heavier than the tail of the t-distribution with $\nu = 2.984$.*
- ❑ The parametric estimates based on the *t*-distribution are similar to the estimates assuming a polynomial tail except when α is very small.
- ❑ In the range 0.01 to 0.2, then VaR is relatively insensitive to the choice of model, except for the poorly fitting normal model.

- ❑ 7_VaR_TailIndex.R



Horizon

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 $\hat{\mu}_P^{1day}$ and $\hat{\sigma}_P^{1day}$ 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

$$\hat{\mu}_P^{M\,days} = M\hat{\mu}_P^{1\,day} \quad \text{and} \quad \hat{\sigma}_P^{M\,day} = \sqrt{M}\hat{\sigma}_P^{1day}$$

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

$$\widehat{VaR_P}^{M\,days} = -S \times \{M\hat{\mu}_p^{1\,day} + \sqrt{M}\hat{\sigma}_P^{1\,day}\Phi^{-1}(\alpha)\}$$

Danger: it assumes normally distributed returns and no autocorrelation or GARCH effects (volatility clustering) of the daily returns.

- ❑ If there is positive autocorrelation, then it underestimates the *M*-day VaR.
- ❑ If the returns are not normally distributed, then there is no simple analog.

Coherent risk measure

Artzner, Delbaen, Eber, and Heath (1997, 1999):
What properties can reasonably be required of a risk measure?

A risk measure R is said to be coherent if it satisfies the following properties:

- Normalized: $R(0) = 0$, the risk of holding no assets is zero.
- Monotonicity: if $\mathrm{P}_1 \leq P_2$, then $R(P_1) \geq R(P_2)$. That is, portfolio P_2 with better values should be less risky than Portfolio P_1 under almost all scenarios.
- Sub-additivity: $R(P_1 + P_2) \leq R(P_1) + R(P_2)$.The risk of two portfolios together cannot get any worse than adding the two risks separately (diversification principle).
- Positive homogeneity: If $c \geq 0$, $R(cP) = cR(p)$. Loosely speaking, if you double your portfolio P then you double your risk.
- Translation invariance: If A is a deterministic portfolio with guaranteed return a $R(P + A) = R(P) - a$ as the portfolio A is just adding cash to your portfolio.

VaR is a wrong risk measure

The VaR bible is the book by Philippe Jorion (2001).The following “definition" is very common:

“VaR is the maximum expected loss of a portfolio over a given time horizon with a certain confidence level."

It is however mathematically meaningless and potentially misleading. In no sense is VaR a maximum loss! We can lose more, sometimes much more, depending on the heaviness of the tail of the loss distribution.

$VaR_{1\%}$ = the minimal loss in the 1%“bad” cases

What about the expected loss in the 1%“bad” cases? → ES

A serious problem with VaR is that it may *discourage* diversification. VaR is not sub-additive, i.e., the desirable property that the risk of an aggregated portfolio is smaller than the sum of the risks of its components may be violated.

- ❑ VaR is an inappropriate risk measure for allocating capital charges – interpreted as trading limits – among organizational units of a bank.
- ❑ VaR is inconsistent with diversification and can thus lead to sub-optimal risk management if used in the context of portfolio optimization or hedging.
- ❑ VaR is inappropriate for the measurement of capital adequacy, as it controls only the probability of default, but not the average loss in the case of default.

Example: VaR is not subadditive

A company is selling par \$1000 bonds with a maturity of one year that pay a simple interest of 5%:

- 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. We assume that the loss is *N*(-50, 1) with probability 0.96 and *N*(1000, 1) with probability 0.04.

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
  - Portfolio 2 buys one bond from each of the two companies.

Both portfolios have the same expected loss, but the second is more diversified.

The loss CDF is:

Portfolio 1:$0.04\Phi(x;\ 2000, 4)\ +\ 0.96\Phi(x; -100, 4)$

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

Example: VaR is not subadditive

- $VaR_{0.05}$ 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.
- VaR depends heavily on the values of α. When α is below the default probability, 0.04, portfolio 1 is more risky than portfolio 2.

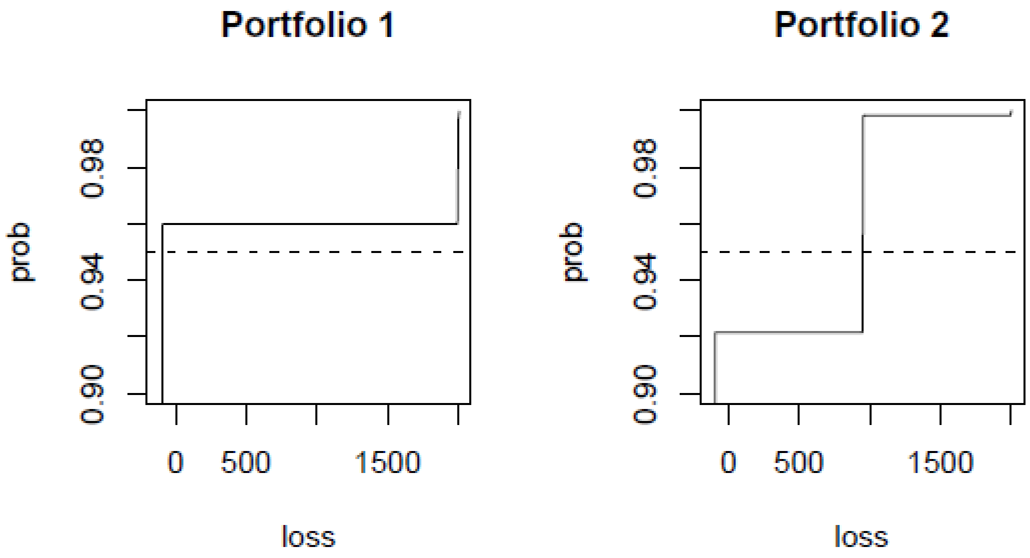
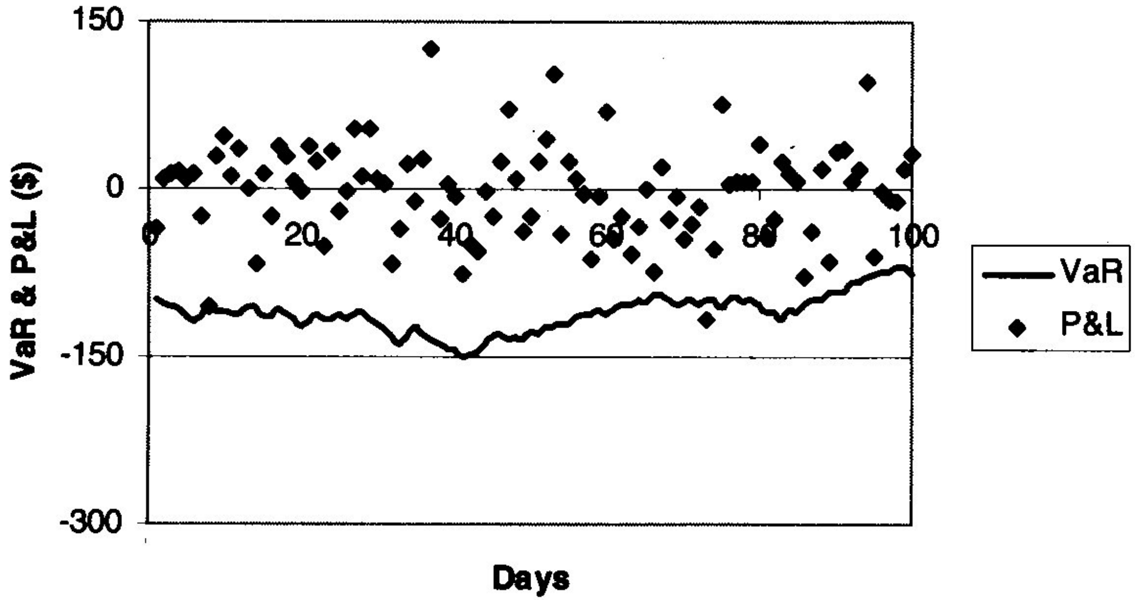


Fig. 19.7. *Example where VaR discourages diversification. Plots of the CDF of the loss distribution. VaR(0.05) is the loss at which the CDF crosses the horizontal dashed line at 0.95.*

- *7_VaR_coherent.R*

Backtesting

Backtesting is to make sure that the probability distribution (e.g., the VaR) is consistent with actual losses. It compares the loss on any given day with the VaR predicted for that day.



Let $I_t = 1\{L_{t+1} > \widehat{VaR}_{t+1|t,\alpha}\}$ denote the indicator for a violation of the predicted VaR for day t + 1.

The exception is Binomial random variable. Binomial variables are those that can have a value of zero or one. It is roughly Bernoulli distributed. The Bernoulli distribution describes the probability of having a given number of outcomes that are equal to one if a binomial variable is sampled multiple times

R lab

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. 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 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. Find $VaR_{0.05}$ and $ES_{0.05}$ for this portfolio.

Rlab7.R

R lab

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.

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.

R Lab

Consider 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. Suppose that $1000 is invested in BMW stock.)

Problem 5. Compute nonparametric estimates of VaR(0.01, 24 hours) and ES(0.01,24 hours).

Problem 6 Compute parametric estimates of VaR(0.01, 24 hours) and ES(0.01,24 hours) assuming that the returns are normally distributed.

Problem 7 Compute parametric estimates of VaR(0.01, 24 hours) and ES(0.01,24 hours) assuming that the returns are t-distributed.

Problem 8 Compare the estimates in 6,7,8. Which do you feel are most realistic?

R lab – Results & Discussions

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. 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.

The R code to fit the multivariate-t model is as Rlab7.R. The estimates are:

```
> muhat
[1]0.0054245 0.0216433
> dfhat
[1]4.2980
> scaleMatrixhat
[,1]         [,2]
[1,]0.0099871 0.0045443
[2,]0.0045443 0.0054012
```

R lab – Results & Discussions

Problem 2 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. Find $VaR_{0.05}$ and $ES_{0.05}$ for this portfolio.

The distribution of the return on the portfolio has a univariate t-distribution. The estimated degrees of freedom is the same as in the multivariate tdistribution for the two stocks, 4.3.

```
alpha = .05
w =  c ( 0.3, 0.7)
muP = w %*% muhat
scaleP = sqrt(w %*% scaleMatrixhat %*% w)
VaR05 =100000*( -muP + scaleP*qt(1-alpha,dfhat))
ES05 = 100000*(-muP + scaleP* dt(qt(alpha,dfhat),dfhat)/alpha*(dfhat + qt(alpha,dfhat)^2)/(dfhat-1) )
muP
scaleP
VaR05
ES05
```

The output contains the estimates of the mean and scale parameter for the return on the portfolio: muP=0.016778, scaleP=0.073851. The estimate of VaR05 is \$13,982 and the estimate of ES05 is \$21,127.

R lab – Results & Discussions

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.

The output shows that the bootstrap confidence interval is $11,900 to $15,638.

R lab – Results & Discussions

Problem 3 Also, plot kernel density estimates of the bootstrap distribution of DF and VaRt (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.

The kernel density estimates are plotted above. Neither appears Gaussian. The density of DF is right-skewed. The skewness of VaR(0.05) is difficult to summarize; there are a few outliers on the right, but if they are ig1wred then the rest of the data look slightly left-skewed. In any case, the shape does not look Gaussian. The Shapiro-Wille tests have very small p-values, which also suggests non-normality.

```
>  shapiro.test(dfhatboot)
Shapiro-Wilk normality test data: dfhatboot
W = 0.841,  p-value  = 2.628e-15

>  shapiro.test(VaR05boot)
Shapiro-Wilk normality test data: VaR05boot
W = 0.6632, p-value < 2.2e-16
```

R lab – Results & Discussions

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.

```
DEC=   Berndt[,2]
Y   =  sort(DEC)
ahat=rep(0,40)
for  (k   in 1:40)
{
ahat[k] =  k/ (sum(log(Y[1:k]/Y[k])))
}
plot ( (1:k) ,ahat ,xlab="n(c)" ,ylab="ahat (c)" ,main="Hill plot")
```

The Hill plot shows a region of stability when n_c is between 5 and 30. Any of the values of a in this region could be used, or one can use the average of all of them.

```
> mean(ahat[5:20])
```

This give us a= 2.15 as one possible answer.

R lab – Results & Discussions

Consider 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. Suppose that $1000 is invested in BMW stock.)

Problem 5. Compute nonparametric estimates of VaR(0.01, 24 hours) and ES(0.01,24 hours).

```
> library(fEcofin)
> data(bmwRet)
> r = bmwRet[,2)
> varNonp =  -1000*quantile(r,.01)
> ind =  (-1000*r > var)
> esNonp =  -1000*sum(r[ind])/ sum(ind)
> varNonp
1%
40.798
> esNonp
[1] 56.492
```

R lab – Results & Discussions

Problem 6 Compute parametric estimates of VaR(0.01, 24 hours) and ES(0.01, 24 hours) assuming that the returns are normally distributed.

```
> varNormal = -1000*qnorm(.01,mean=mean(r),sd=sd(r))
> ESNormal = 1000*( -mean(r) + sd(r)*dnorm(qnorm(.01))/.01
> varNormal
[1]33.986
> ESNormal
[1]38.986
```

R lab – Results & Discussions

Problem 7 Compute parametric estimates of VaR(0.01, 24 hours) and ES(0.01, 24 hours) assuming that the returns are t-distributed.

```
> loglik = function(theta){
+ x =  (r-theta[1])/theta[2]
+ - sum(log(dt(x,theta[3])/theta[2]))
+ }
> theta0 = c(mean(r),sd(r),5)
> fit = optim(theta0,loglik)
> theta = fit$par
> theta
[1]0.00013196 0.00926451 2.98858766
> alpha = .01
> VaRt =1000*(-theta[1] + theta[2]*qt(1-alpha,theta[3]))
>ESt= 1000*(-theta[1] + theta[2]* dt(qt(alpha,theta[3]),theta[3])/alpha*
+    (theta[3] + qt(alpha,theta[3])-2)/(theta[3]-1))
> VaRt
[1]42.064
> ESt
[1]65.058
```

R lab – Results & Discussions

Problem 8 Compare the estimates in 6,7,8. Which do you feel are most realistic?

The three sets of estimated risk measures are below.

Method	VaR	ES
Nonparametric	40.8	56.5
Normal	34.0	39.0
t	42.1	65.1

To see which might be most realistic, a normal plot of the returns is shown below. The returns have much heavier tails than a normal distribution, so the risk measures based on normality arc not realistic. There are 6146 returns, so the nonparametric estimates are reasonable, especially VaR. However, even 6146 returns might not be enough to estimate the tails accurately using nonparametric estimators, so I would prefer the t-based estimates.

We can compare the nonparametric and t-based estimators of VaR and ES by resampling. Both model-based (assuming t-distributed returns) and model-free resan1pling could be used. It would be interesting to compare the two types of estimators using bias, standard deviation, and mean squared error.