

Modelling of Long-Term Risk

Paul Embrechts

ETH Zurich

`embrechts@math.ethz.ch`

Roger Kaufmann

Swiss Life

`roger.kaufmann@swisslife.ch`

©2004 (P. Embrechts and R. Kaufmann)

Contents

- A. Basel II
- B. Scaling of Risks
- C. One-Year Risks
- D. Conclusions and Further Work

A. Basel II

Amendment to the Capital Accord to Incorporate Market Risks
(Basle Committee on Banking Supervision, 1996):

- “In calculating the **value-at-risk**, a **99th percentile**, one-tailed confidence interval is to be used.”
- “In calculating value-at-risk, an instantaneous price shock equivalent to a **10 day** movement in prices is to be used.”
- “Banks may use value-at-risk numbers calculated according to shorter holding periods **scaled up to ten days by the square root of time.**”

Basel II (cont.)

- **Market risk:** 10-day value-at-risk, 99%
Standard: 1-day value-at-risk, 95%
- **Insurance:** 1-year value-at-risk, 99%
1-year expected shortfall, 99%

Value-at-Risk and Expected Shortfall

- Primary risk measure: **Value-at-Risk** defined as

$$\text{VaR}_p(X) = F_{-X}^{-1}(p),$$

i.e. the p th quantile of F_{-X} . (X denotes the profit, $-X$ the loss.)

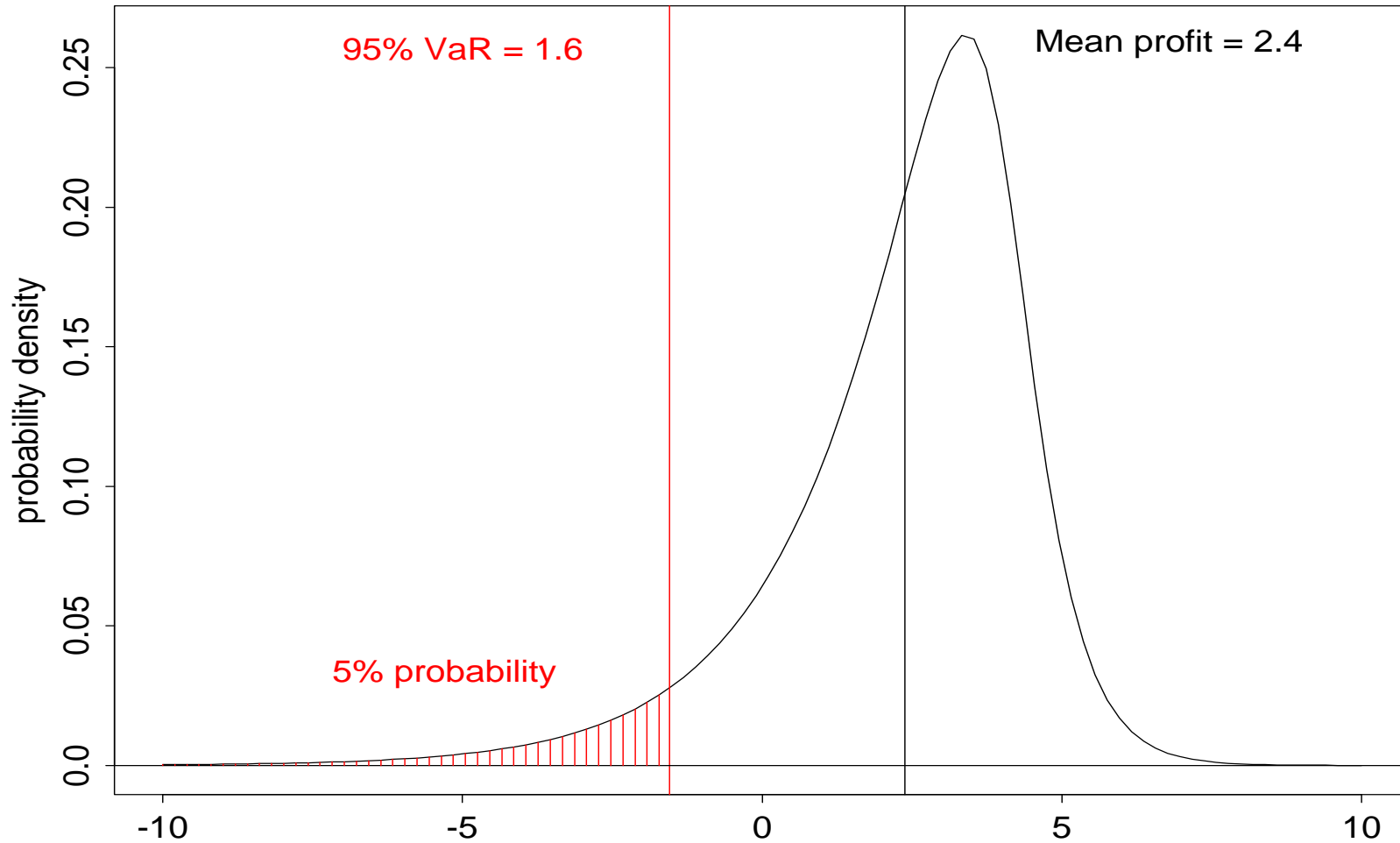
- Alternative risk measure: **Expected shortfall** defined as

$$\text{ES}_p(X) = E(-X \mid X < -\text{VaR}_p);$$

i.e. the **average** loss when VaR is exceeded. $S_p(X)$ gives information about **frequency and size** of large losses.

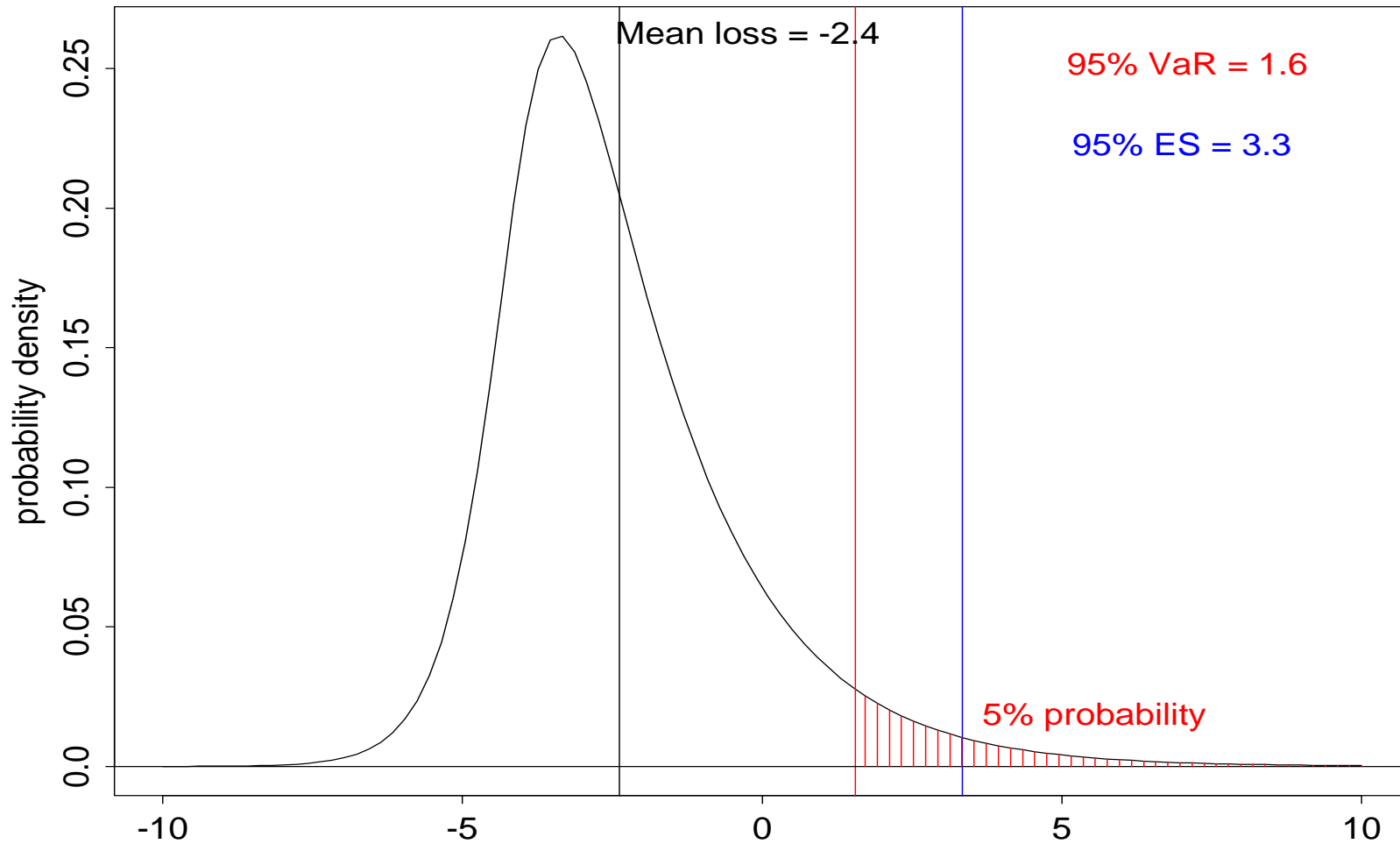
VaR in Visual Terms

Profit & Loss Distribution (P&L)



Losses and Profits

Loss Distribution



B. Scaling

Question 1: How to get a 10-day VaR (or 1-year VaR)?

Solution in the praxis: scale the 1-day VaR by $\sqrt{10}$ (or $\sqrt{250}$).

Question 2: How good is scaling?

→ Model dependent!

Scaling under Normality

Under the assumption

$$X_i \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2),$$

n -day log-returns are normally distributed as well:

$$\sum_{i=1}^n X_i \sim \mathcal{N}(0, n\sigma^2).$$

For a $\mathcal{N}(0, \tilde{\sigma}^2)$ -distributed profit X , $\text{VaR}_p(X) = \tilde{\sigma} x_p$, where x_p denotes the p -Quantile of a standard normal distribution. Hence

$$\text{VaR}^{(n)} = \sqrt{n} \text{VaR}^{(1)}.$$

Accounting for Trends

When adding a constant trend μ ,

$$X_i \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2),$$

n -day log-returns are still normally distributed:

$$\sum_{i=1}^n X_i \sim \mathcal{N}(n\mu, n\sigma^2).$$

Hence

$$\text{VaR}^{(n)} + n\mu = \sqrt{n} (\text{VaR}^{(1)} + \mu),$$

i.e.

$$\text{VaR}^{(n)} = \sqrt{n} \text{VaR}^{(1)} - (n - \sqrt{n})\mu.$$

Autoregressive Models

For an autoregressive model of order 1,

$$X_t = \lambda X_{t-1} + \epsilon_t, \quad \epsilon_t \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2),$$

1-day and n -day log-returns are normally distributed:

$$X_t \sim \mathcal{N}\left(0, \frac{\sigma^2}{1 - \lambda^2}\right)$$

and

$$\sum_{i=1}^n X_i \sim \mathcal{N}\left(0, \frac{\sigma^2}{(1 - \lambda)^2} \left(n - 2\lambda \frac{1 - \lambda^n}{1 - \lambda^2}\right)\right).$$

Scaling for AR(1) Models

For an AR(1) model with normal innovations,

$$\frac{\text{VaR}^{(n)}}{\text{VaR}^{(1)}} = \sqrt{\frac{1 + \lambda}{1 - \lambda} \left(n - 2\lambda \frac{1 - \lambda^n}{1 - \lambda^2} \right)}.$$

For small values of λ , $\sqrt{n} \text{VaR}^{(1)}$ is a good approximation of $\text{VaR}^{(n)}$.

Non-Normal Innovations

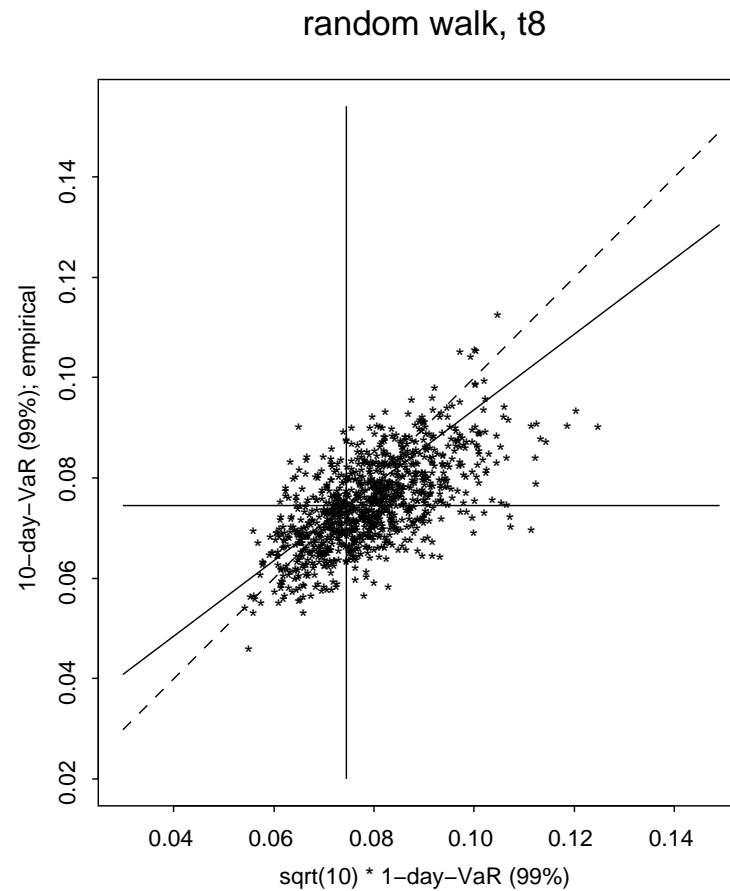
Question: Is scaling with \sqrt{n} still appropriate if innovations are non-normal?

Example: random walk, $X_i \stackrel{\text{i.i.d.}}{\sim} t_8$

Based on 250 log-returns, how good is $\sqrt{10} \cdot \widehat{\text{VaR}}_{99\%}^{(1)}$ as an estimate for the 10-day 99% VaR?

($\widehat{\text{VaR}}_{99\%}^{(1)}$ denotes the one-day 99% VaR estimate.)

Non-Normal Innovations (cont.)



Scaling is still good, but other methods like random resampling perform slightly better.

AR(1)-GARCH(1,1) Processes

A more complex process, often used for practical applications, is the GARCH(1,1) process ($\lambda = 0$) and its generalization, the AR(1)-GARCH(1,1) process:

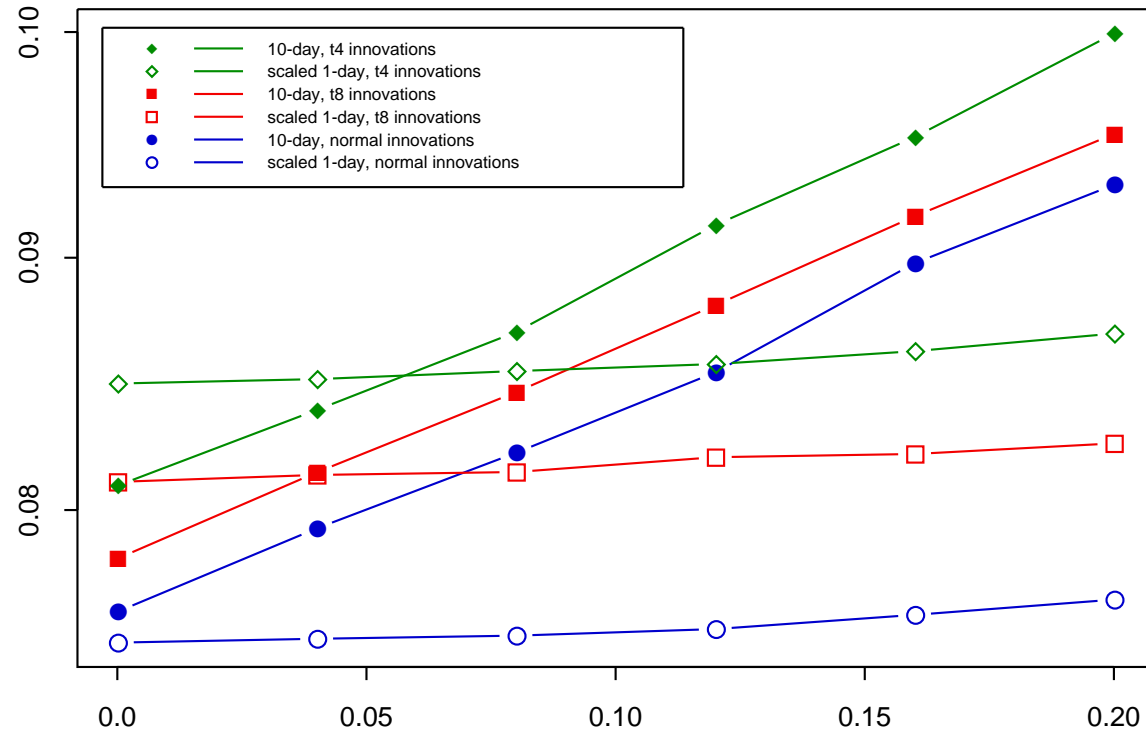
$$X_t = \lambda X_{t-1} + \sigma_t \epsilon_t,$$

$$\sigma_t^2 = a_0 + a(X_{t-1} - \lambda X_{t-2})^2 + b \sigma_{t-1}^2,$$

$$\epsilon_t \text{ i.i.d.}, E[\epsilon_t] = 0, E[\epsilon_t^2] = 1.$$

(typical parameters: $\lambda = 0.04$, $a_0 = 3 \cdot 10^{-6}$, $a = 0.05$, $b = 0.92$)

Scaling for AR(1)-GARCH(1,1) Processes



Goodness of fit of the scaling rule, depending on different values of λ (x axis) for different distributions of the innovations ϵ_t .

For typical parameters ($\lambda = 0.04$, $\epsilon_t \sim t_8$), the fit is almost perfect.

GARCH(1,1) vs. Random Walk

A GARCH(1,1) process

$$X_{a,t} = \sigma_{a,t} \epsilon_t,$$

$$\sigma_{a,t}^2 = a_0 + a X_{a,t-1}^2 + b \sigma_{a,t-1}^2,$$

$$\epsilon_t \text{ i.i.d.}, E[\epsilon_t] = 0, E[\epsilon_t^2] = 1,$$

(where a is typically close to 0) can be approximated by a process with variance

$$\sigma_{0,t}^2 = a_0 + b \sigma_{0,t-1}^2$$

or

$$\sigma_{0,t}^2 = a_0 + (a + b) \sigma_{0,t-1}^2.$$

GARCH(1,1) vs. Random Walk (cont.)

If the initial values of the processes $(X_{a,t})$ and $(X_{0,t})$ coincide, then

$$E[(X_{a,t} - X_{0,t})^2] \leq \text{fct(parameters)},$$

and

$$E\left[\left(\sum_{t=n+1}^{n+h} X_{a,t} - \sum_{t=n+1}^{n+h} X_{0,t}\right)^2\right] \leq \text{fct(parameters)}.$$

These inequalities can be used to get bounds for (conditional and unconditional) value-at-risk of GARCH(1,1) processes. Analogously, value-at-risk estimates for AR(1)-GARCH(1,1) processes can be obtained by approximating them with AR(1) processes.

Stochastic Volatility Model with Jumps

An alternative to autoregressive types of models are stochastic volatility models:

$$X_t = a \sigma_t Z_t + b J_t \epsilon_t,$$

$$\sigma_t = \sigma_{t-1}^\phi e^{c Y_t},$$

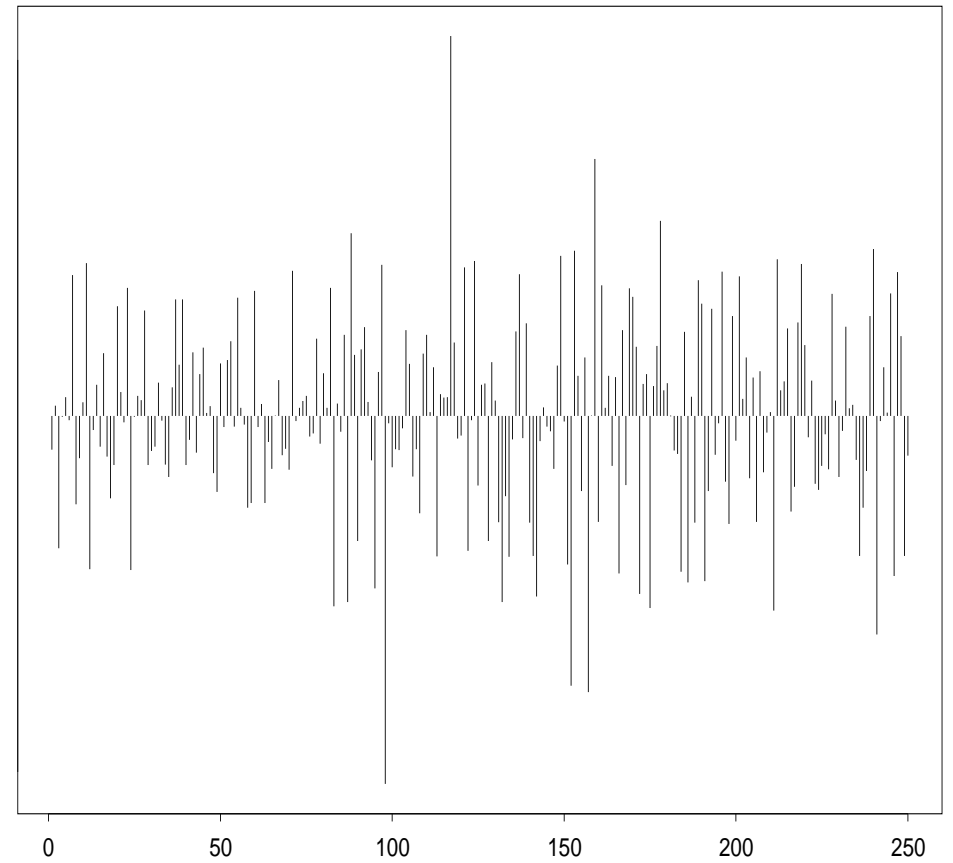
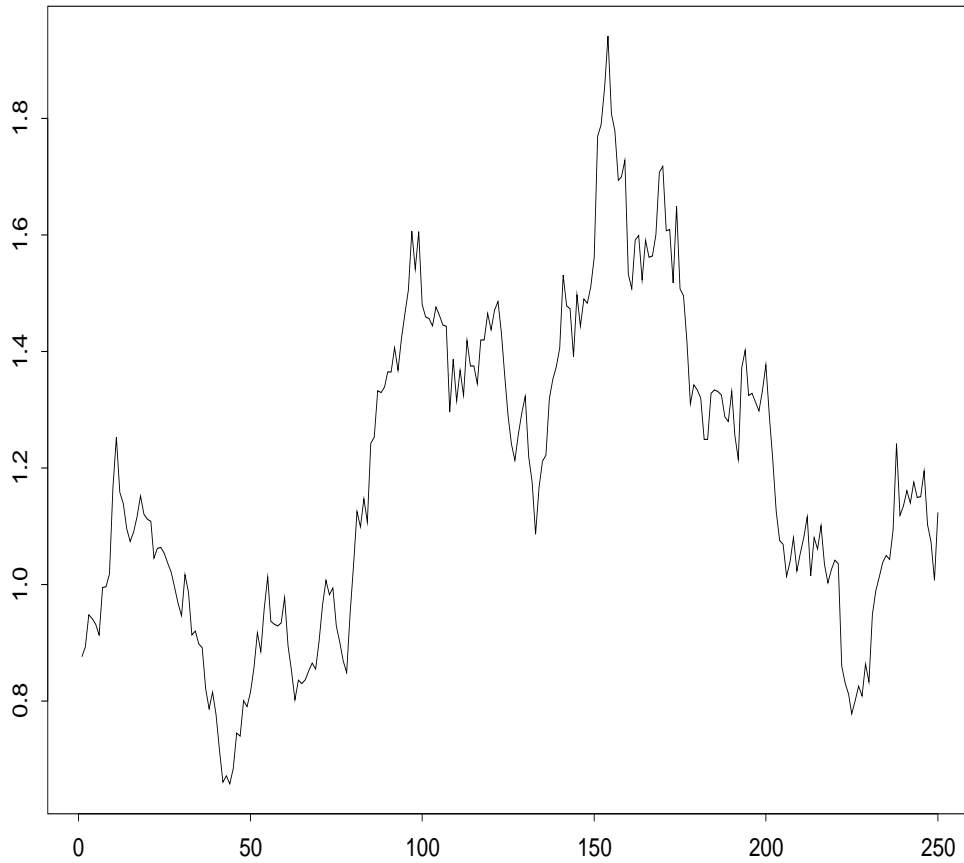
$$\epsilon_t, Z_t, Y_t \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1),$$

$$J_t \stackrel{\text{i.i.d.}}{\sim} \text{Bernoulli}(\lambda)$$

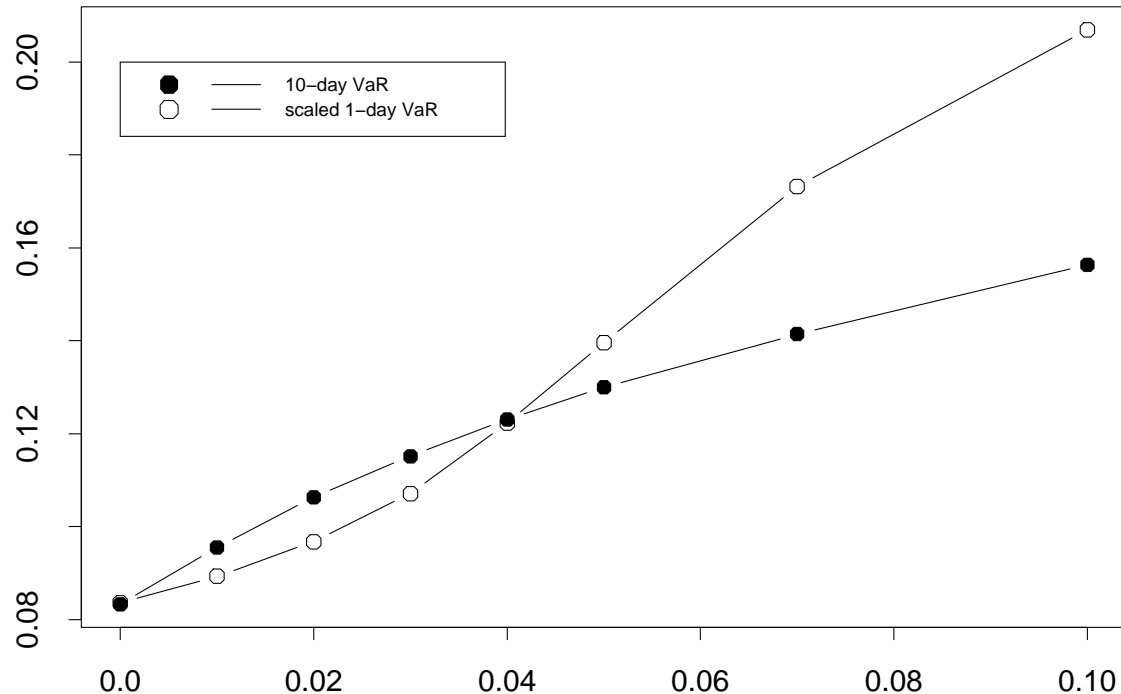
(typical parameters:

$$\lambda = 0.01, a = 0.01, b = 0.05, c = 0.05, \phi = 0.98)$$

Stochastic Volatility Model: Volatility and Returns



Scaling in the Stochastic Volatility Model



Goodness of fit of the scaling rule, depending on different values of λ (x axis).

The scaled 1-day VaR underestimates the 10-day VaR for small values of λ . For $\lambda > 0.04$, this changes to an overestimation.

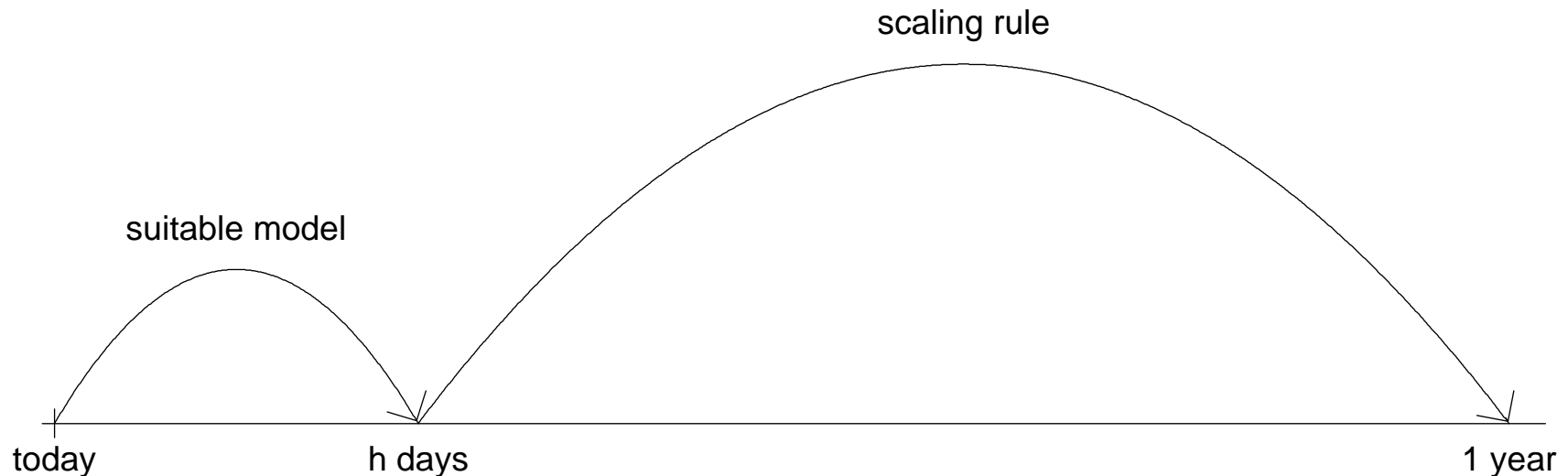
C. One-Year Risks

Problems when modelling yearly data:

- Non-stationarity of data sets.
- Lack of yearly returns.
- Properties of yearly data are different from those of daily data.

How to Estimate Yearly Risks

- Fix a horizon $h < 1$ year, for which data can be modelled.
- Use a scaling rule for the gap between h and 1 year.



Models

- Random Walks
- Autoregressive Processes
- GARCH(1,1) Processes
- Heavy-tailed Distributions

Random Walk

Financial log-data $(s_t)_{t \in h\mathbb{N}}$ can be modelled as a random walk process with constant trend and normal innovations:

$$s_t = s_{t-h} + X_t, \quad X_t \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2) \quad \text{for } t \in h\mathbb{N}.$$

The square-root-of-time rule (accounting for the trend) can be used to scale h -day risks to 1-year risks.

Autoregressive Processes

For an AR(p) model with trend and normal innovations,

$$s_t = \sum_{i=1}^p a_i s_{t-ih} + \epsilon_t \quad \text{for } t \in h\mathbb{N},$$

$$(\epsilon_t \sim \mathcal{N}(\mu_0 + \mu_1 t, \sigma^2), \text{ independent})$$

the 1-year value-at-risk and expected shortfall can be calculated as a function of the parameters μ_1 , σ and a_i , and the current and past values of (s_t) .

Generalized Autoregressive Conditional Heteroskedastic Processes

Assuming a GARCH(1,1) process with Student- t distributed innovations for h -day log-returns,

$$X_t = \mu + \sigma_t \epsilon_t \quad \text{for } t \in h\mathbb{N},$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 (X_{t-h} - \mu)^2 + \beta_1 \sigma_{t-h}^2,$$

where $\epsilon_t \stackrel{\text{i.i.d.}}{\sim} t_\nu$, $E[\epsilon_t] = 0$, $E[\epsilon_t^2] = 1$,

1-year log-returns follow a so-called weak GARCH(1,1) process. The corresponding VaR and ES can be calculated as a function of the above parameters and the current and past values of (X_t) .

Heavy-tailed Distributions

h -day log-returns $(X_t)_{t \in h\mathbb{N}}$ are said to have a heavy-tailed distribution, if

$$P[X_t < -x] = x^{-\alpha} L(x) \quad \text{as } x \rightarrow \infty,$$

where $\alpha \in \mathbb{R}^+$ and L is a slowly varying function, i.e. $\lim_{x \rightarrow \infty} \frac{L(sx)}{L(x)} = 1$ for all $s > 0$.

Also in this case, 1-year VaR and ES can be estimated based on the parameter α and on the observed data.

Backtesting

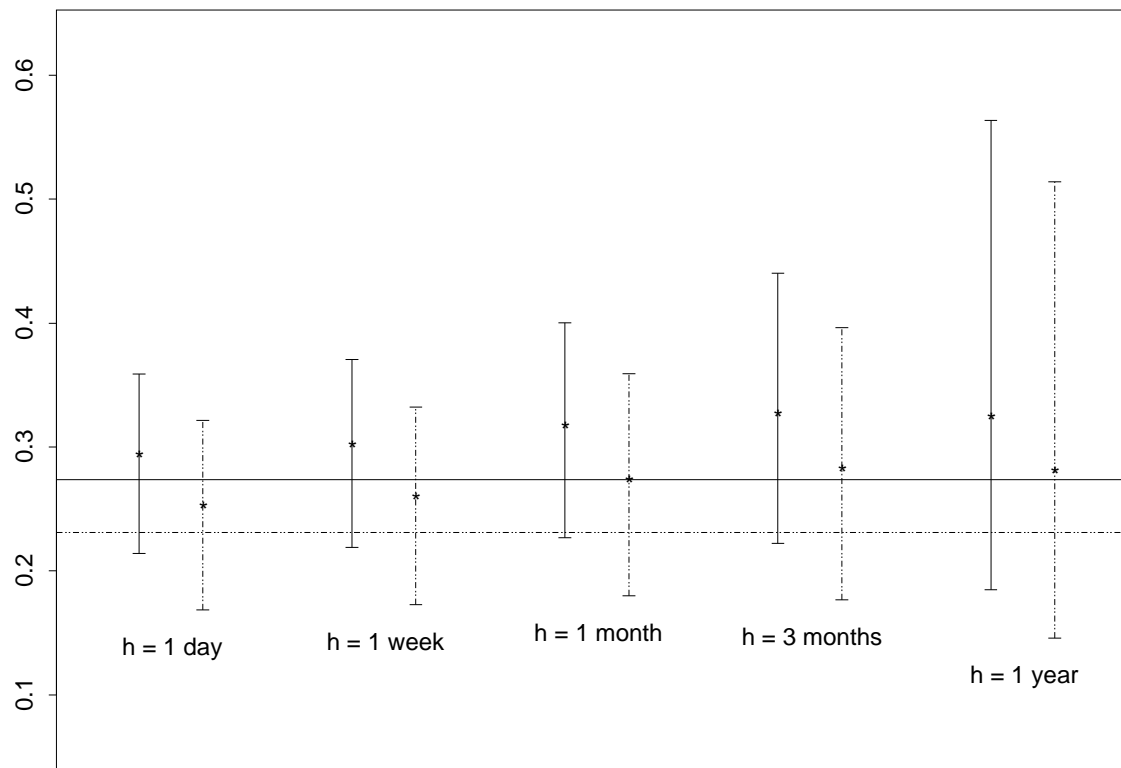
The suitability of these models for estimating one-year financial risks can be assessed by comparing estimated value-at-risk and expected shortfall with observed return data for

- stock indices,
- foreign exchange rates,
- 10-year government bonds,
- single stocks.

Conclusions for 1-Year Forecasts

- The **random walk** model performs in general better than the other models under investigation. It provides **satisfactory results across all classes of data** and for both confidence levels investigated (95%, 99%). However, like all the other models under investigation, the risk estimates for single stocks are not as good as those for foreign exchange rates, stock indices, and 10-year bonds.
- The **optimal calibration horizon** is about **one month**. Based on these data, the square-root-of-time rule (accounting for trends) can be applied for estimating one-year risks.

Confidence Intervals for a Random Walk



Point estimates and 95% confidence intervals for one-year 99% expected shortfall and 99% value-at-risk (percentage loss) for a simulated random walk with normal innovations.

D. Conclusions

- The square-root-of-time scaling rule performs very well to scale risks from a short horizon (1 day) to a longer one (10 days, 1 year).
- The reasons for this good performance are non-trivial. Each situation has to be investigated separately. The square-root-of-time rule should not be applied before checking its appropriateness.
- In the limit, as $\alpha \rightarrow 1$, scaling a short-term VaR_α to a long-term risk using the square-root-of-time rule is for most situations not appropriate any more.

Further Work

- An interesting subject for further research is to find the limits, where the square-root-of-time rule fails. For example changing one single parameter in a model can have a strong effect on the appropriateness of this scaling rule.
- Linked to this topic is the model-dependent question, *why* the square-root-of-time rule performs well (or not so well) in a certain situation.
- An interesting generalisation of this work would be the investigation of multivariate models.

Bibliography

- [Brummelhuis and Kaufmann, 2004a] Brummelhuis, R. and Kaufmann, R. (2004a). *GARCH(1,1) and AR(1)-GARCH(1,1) processes: perturbation estimates for value-at-risk*. Working Paper.
- [Brummelhuis and Kaufmann, 2004b] Brummelhuis, R. and Kaufmann, R. (2004b). *Time Scaling for GARCH(1,1) and AR(1)-GARCH(1,1) Processes*. Working Paper.
- [Embrechts et al., 2004] Embrechts, P., Kaufmann, R., and Patie, P. (2004). Strategic long-term financial risks: Single risk factors. *To appear in: A special issue of Computational Optimization and Applications*.

[Kaufmann, 2004a] Kaufmann, R. (2004a). *Long-Term Risk Management*. PhD Thesis, ETH Zurich.

[Kaufmann, 2004b] Kaufmann, R. (2004b). *Time Scaling for Stochastic Volatility Models*. Working Paper, ETH Zurich.

[Kaufmann and Patie, 2003] Kaufmann, R. and Patie, P. (2003). *Strategic Long-Term Financial Risks: The One Dimensional Case*. RiskLab Report, ETH Zurich. Available at <http://www.risklab.ch/Papers.html#SLTFR>.