Backtesting Value-at-Risk Models

Christophe Hurlin

University of Orléans

Séminaire Validation des Modèles Financiers. 29 Avril 2013

Introduction

The Value-at-Risk (VaR) and more generally the Distortion Risk Measures (Expected Shortfall, etc.) are standard risk measures used in the current regulations introduced in Finance (Basel 2), or Insurance (Solvency 2) to fix the required capital (Pillar 1), or to monitor the risk by means of internal risk models (Pillar 2).

Introduction

Definition

Let $\{r_t\}_{t=1}^{T}$ be a given P&L series. The daily (conditional) VaR for a nominal coverage rate α is defined as

$$\Pr\left[r_t < -VaR_{t|t-1}(\alpha) \mid \mathcal{F}_{t-1} \right] = \alpha$$

where $\mathcal{F}_{t-1}$ denotes the set of information available at time $t-1$.

Introduction

Who does use VaR?	What for?
Bank risk manager	Measure firm-level market, credit, op. risk
Bank executives	Set limits (management)
Banking regulators	Determine capital requirements
Exchanges	Compute margins
Regulators	Forecast systemic risk (CoVaR)
Industry	Ex: EDF, spot prices of electricity

"Disclosure of quantitative measures of market risk, such as value-at-risk, is enligthening only when accompanied by a thorough discussion of how the risk measures were calculated and how they related to actual performance", Alan Greenspan (1996)

Introduction

Definition

Backtesting is a set of statistical procedures designed to check if the real losses are in line with VaR forecasts (Jorion, 2007).

Introduction

- Whatever the type of use of VaR, the VaR forecasts are generated by an internal risk model.
- This model is used to produced a sequence of pseudo out-of sample VaR forecasts for a past period (typically one year)
- The backtesting is based on the comparison of the observed P&L to these VaR forecasts.

Outlines

1. How to test the validity of a VaR model?
2. What are the backtesting strategies?
3. What are the good practices?

Backtesting Principles

Backtesting Principles

Remark 1: Ex-post VaR is not observable, so it is impossible to compute traditional statistics or criteria such as MSFE.

Remark 2: There is no proxy for the VaR contrary to the volatility (realized volatility, Andersen and Bollerslev 1998)

Patton, A.J. (2011) Volatility forecast comparison using imperfect volatility proxies, Journal of Econometrics, 260, 246-256.

Backtesting Principles

Backtesting procedures are based on VaR exceptions

Definition

We denote $I_t(\alpha)$ the hit variable associated to the *ex-post* observation of an $\alpha\%$ VaR exception at time t :

$$I_t(\alpha) = \begin{cases} 1 & \text{if } r_t < -VaR_{t|t-1}(\alpha) \\ 0 & \text{else} \end{cases}$$

Backtesting Principles

Christoffersen (1998) : VaR forecasts are valid if and only if the violation process $I_t(\alpha)$ satisfies the following two assumptions:

1. The unconditional coverage (UC) hypothesis.
2. The independence (IND) hypothesis.

Christoffersen P.F. (1998), Evaluating interval forecasts, International Economic Review, 39, pp. 841-862.

Backtesting Principles

Definition (unconditional coverage hypothesis)

The unconditionnal probability of a violation must be equal to the α coverage rate

$$\Pr\left[I_t(\alpha)=1\right]=\mathbb{E}\left[I_t(\alpha)\right]=\alpha.$$

- If $\Pr\left[I_t(\alpha)=1\right]>\alpha$, the risk is under-estimated
- If $\Pr\left[I_t(\alpha)=1\right]<\alpha$, the risk is over-estimated

Backtesting Principles

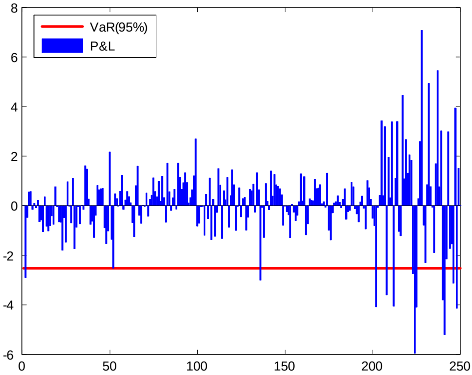
Definition (independence hypothesis)

VaR violations observed at two different dates must be independently distributed.

$$I_t(\alpha) \text{ and } I_s(\alpha) \text{ are independently distributed for } t \neq s$$

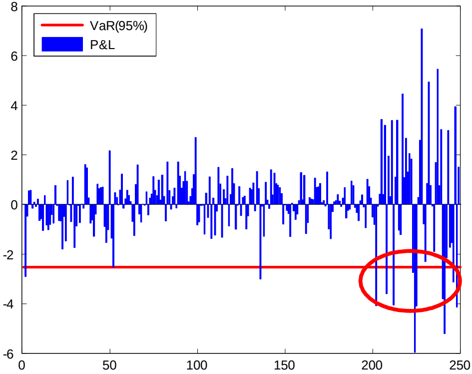
Backtesting Principles

Figure: Illustration: violations' cluster



Backtesting Principles

Figure: Illustration: violations' cluster



Backtesting Principles

Definition (conditional coverage hypothesis)

The violation process satisfies a difference martingale assumption.

$$\mathbb{E}\left[\, I_t(\alpha) \mid \mathcal{F}_{t-1} \right] = \alpha$$

Backtesting Principles

Remark: These assumptions can be expressed as distributional assumptions.

Under the UC assumption, each variable $I_t(\alpha)$ has a Bernouilli distribution with a probability α.

$$I_{tt}(\alpha) \sim Bernouilli(\alpha)$$

Under the IND assumption, these variables are independent, and the number of violations has a Binomial distribution.

$$\sum_{t=1}^{T} I_t(\alpha) \sim B(T, \alpha)$$

Testing strategies

What are the backtesting strategies?

Testing strategies

Let us consider a sequence of daily VaR out-of-sample forecasts $\left\{VaR_{t|t-1}(\alpha)\right\}_{t=1}^{T}$ and the corresponding observed P&L.

How to test the validity of the internal risk model?

Hurlin C. and Pérignon C. (2012), Margin Backtesting, Review of Futures Market, 20, pp. 179-194

Testing strategies

Testing strategies:

1. Frequency-based tests
2. Magnitude-based tests
3. Multivariate tests
4. Independence tests
5. Duration-based tests

Testing strategies: frequency-based tests (1/5)

Figure: BIS "Traffic Light" System

Zone	Number of exceptions	Increase in scaling factor
Green Zone	0	0,00
	1	0,00
	2	0,00
	3	0,00
	4	0,00
Yellow Zone	5	0,40
	6	0,50
	7	0,65
	8	0,75
	9	0,85
Red Zone	10 or more	1,00

Note: VaR(1%, 1 day), 250 daily observations

Testing strategies: frequency-based tests (1/5)

Definition

Christoffersen (1998) proposes a Likelihood Ratio statistic for UC defined as:

$$
\begin{aligned}
LR_{UC} &= -2\ln\left[(1-\alpha)^{T-H}\alpha^{H}\right] \\
&\quad +2\ln\left[(1-H/T)^{T-H}(H/T)^{H}\right] \underset{T\to\infty}{\overset{d}{\longrightarrow}} \chi^{2}(1)
\end{aligned}
$$

where $H=\sum_{t=1}^{T} I_t(\alpha)$ denotes the total number of exceedances.

- For a nominal risk of 5%, the null of UC can not be rejected if and only if $H<7$ for $T=250$ and $\alpha=1\%$.

Testing strategies: frequency-based tests (1/5)

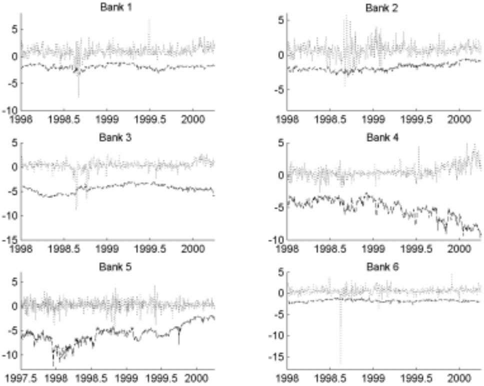
Example

Berkowitz and O-Brien (2002) consider the VaR forecasts of six US commercial banks

Berkowitz, J., and O-Brien J. (2002), How Accurate are the Value-at-Risk Models at Commercial Banks, *Journal of Finance*.

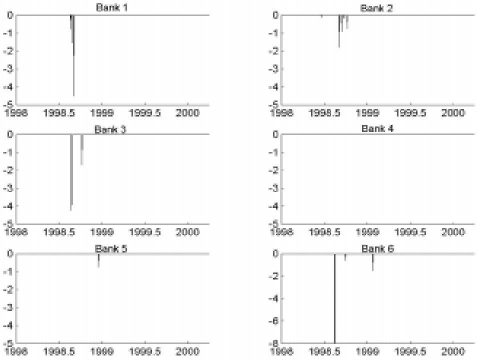
Testing strategies: frequency-based tests (1/5)

Figure: Bank Daily VaR Models



Testing strategies: frequency-based tests (1/5)

Figure: Violations of Banks' 99% VaR



Testing strategies: frequency-based tests (1/5)

Table 4. Backtests of Large-Scale VaR Models

	Violation Rate	Coverage	Conditional Coverage	Independence	Serial Correlation
Bank 1	0.005	1.54 [.214]	1.57 [.455]	.0321 [.858]	-.00533 [.885]
Bank 2	0.010	.00693 [.934]	4.01 [.135]	4.00* [.046]	.158* [.016]
Bank 3	0.005	1.70 [.193]	8.81* [.012]	7.11** [.008]	.330** [.001]
Bank 4	0.000	NA	NA	NA	NA
Bank 5	0.001	8.92** [.003]	8.93* [.012]	.00271 [.959]	-.00137 [.995]
Bank 6	0.005	1.71 [.191]	1.74 [.419]	.0312 [.860]	-.00511 [.885]

Testing strategies: frequency-based tests (1/5)

Table 4. Backtests of Large-Scale VaR Models

	Violation Rate	Coverage	Conditional Coverage	Independence	Serial Correlation
Bank 1	0.005	1.54 [.214]	1.57 [.455]	.0321 [.858]	-.00533 [.885]
Bank 2	0.010	.00693 [.934]	4.01 [.135]	4.00* [.046]	.158* [.016]
Bank 3	0.005	1.70 [.193]	8.81* [.012]	7.11** [.008]	.330** [.001]
Bank 4	0.000	NA	NA	NA	NA
Bank 5	0.001	8.92** [.003]	8.93* [.012]	.00271 [.959]	-.00137 [.995]
Bank 6	0.005	1.71 [.191]	1.74 [.419]	.0312 [.860]	-.00511 [.885]

Testing strategies

Testing strategies:

1. Frequency-based tests
2. Magnitude-based tests
3. Multivariate tests
4. Independence tests
5. Duration-based tests

Testing strategies: magnitude-based tests (2/5)

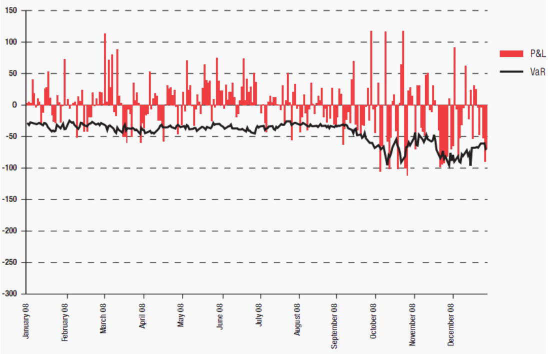
All these tests do not take into account the **magnitude of the losses beyond the VaR**

Example

Consider two banks that both have a one-day 1%-VaR of \$100 million. Assume each bank reports three VaR exceptions, but the average VaR exceedance is \$1 million for bank A and \$500 million for bank B.
In this case, **standard backtesting methodologies** would indicate that the **performance of both models is equal** and **acceptable**.

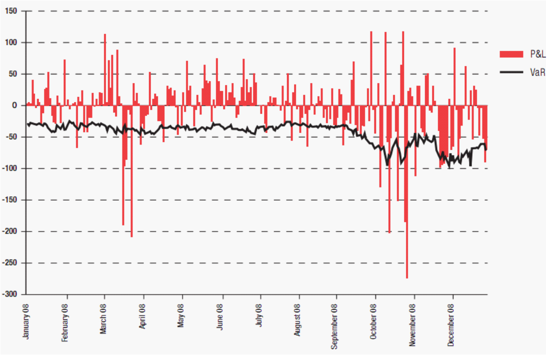
Testing strategies: magnitude-based tests (2/5)

Figure: Daily VaR and P/L for SocGen 2008



Testing strategies: magnitude-based tests (2/5)

Figure: Daily VaR and P/L for SocGen 2008



Testing strategies: magnitude-based tests (2/5)

The Risk Map

Colletaz G., Hurlin C. and Perignon C. (2013), The Risk Map: a new tool for Risk Management, forthcoming in Journal of Banking and Finance

Testing strategies: magnitude-based tests (2/5)

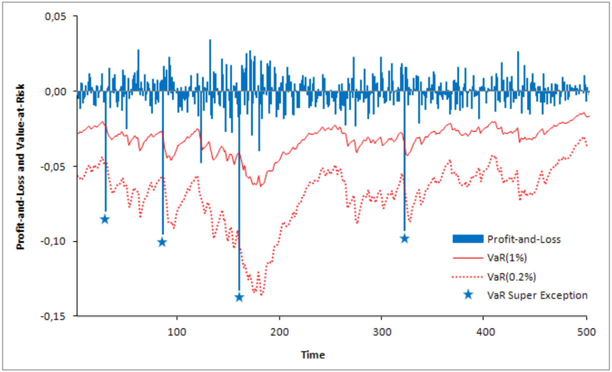
We propose a VaR backtesting methodology based on the number and the severity of VaR exceptions: this approach exploits the concept of "super exception".

Definition

We define a super exception using a VaR with a much smaller coverage probability α', with $\alpha' < \alpha$. In this case, a super exception is defined as a loss greater than $VaR_t(\alpha')$.

Testing strategies: magnitude-based tests (2/5)

Figure: VaR Exception vs. VaR Super Exception



Testing strategies: magnitude-based tests (2/5)

Solution

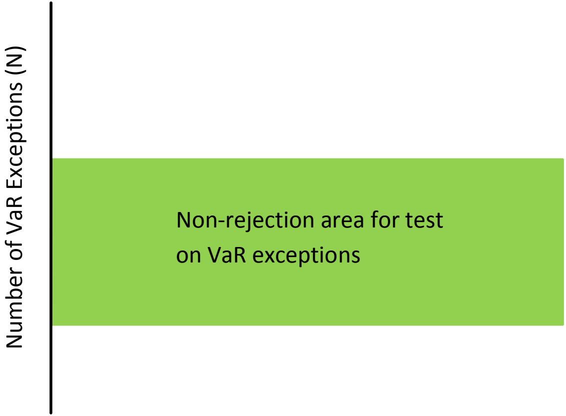
Given VaR exceptions $I_t(\alpha)$ and VaR super exception $I_t(\alpha')$, we define a **Risk Map** *that jointly accounts for the number and the magnitude of the VaR exceptions*

Let us consider a given UC test with a statistic $Z(\alpha)$ based on the violations sequence $\{I_t(\alpha)\}_{t=1}^{T}$.

$$H_0 : \mathbb{E}\left[I_t(\alpha)\right] = \alpha$$

$$H_1 : \mathbb{E}\left[I_t(\alpha)\right] \neq \alpha.$$

Testing strategies: magnitude-based tests (2/5)



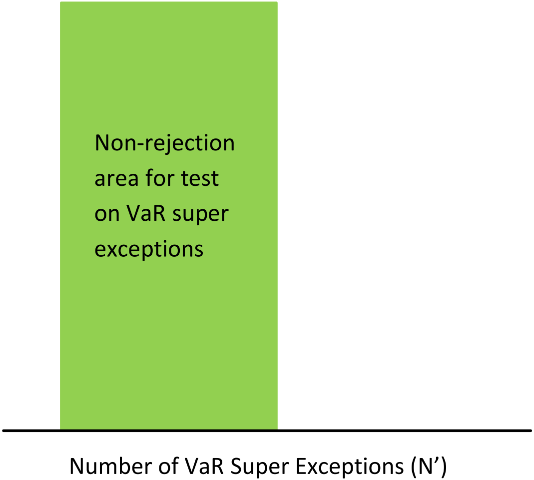
Testing strategies: magnitude-based tests (2/5)

Based on the same UC test, it is possible to test for **the magnitude of VaR exceptions**, via the **VaR super exceptions** $\{I_t(\alpha')\}_{t=1}^{T}$

$$H_0 : \mathbb{E}\left[I_t\left(\alpha'\right)\right] = \alpha'$$

$$H_1 : \mathbb{E}\left[I_t\left(\alpha'\right)\right] \neq \alpha'$$

Testing strategies: magnitude-based tests (2/5)



Testing strategies: magnitude-based tests (2/5)

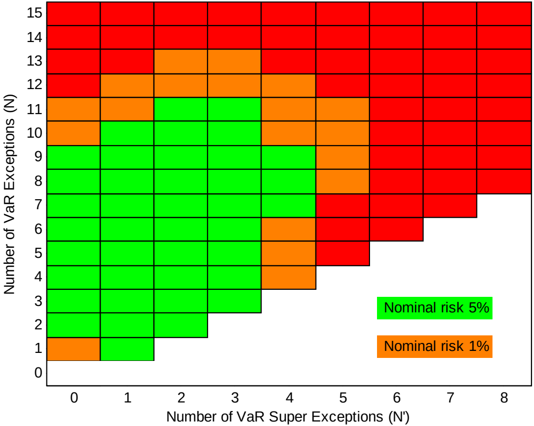
We can also **test jointly** for both magnitude and frequency of VaR exceptions:

$$H_0 : \mathbb{E}\left[I_t\left(\alpha\right)\right] = \alpha \text{ and } \mathbb{E}\left[I_t\left(\alpha'\right)\right] = \alpha'$$

Multivariate approach

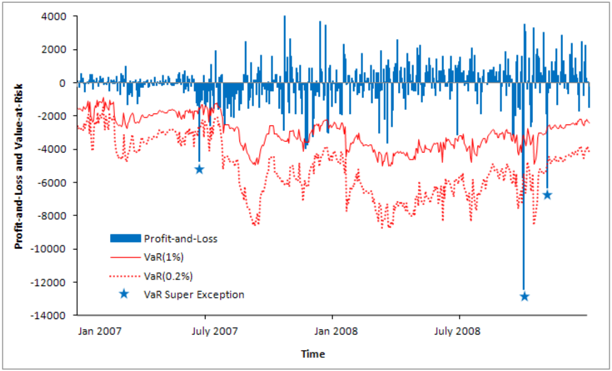
Perignon C. and Smith, D. (2008), A New Approach to Comparing VaR Estimation Methods, Journal of Derivatives

Testing strategies: magnitude-based tests (2/5)

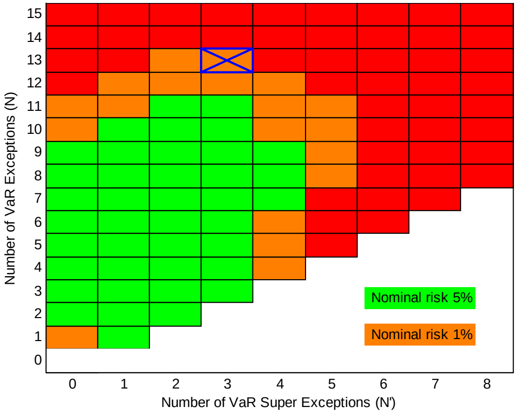


Testing strategies: magnitude-based tests (2/5)

Figure: Backtesting Bank VaR: La Caixa (2007-2008)



Testing strategies: magnitude-based tests (2/5)



Testing strategies

Testing strategies:

1. Frequency-based tests
2. Magnitude-based tests
3. Multivariate tests
4. Independence tests
5. Duration-based tests

Testing strategies: multivariate tests (3/5)

Intuition: Testing the validity of the VaR model for M coverage rates, with $M > 1$.

- **Perignon C. and Smith, D. (2008), A New Approach to Comparing VaR Estimation Methods, Journal of Derivatives**
- **Hurlin C. and Tokpavi, S. (2006), "Backtesting Value-at-Risk Accuracy: A Simple New Test", Journal of Risk**

Testing strategies: multivariate tests (3/5)

Perignon and Smith (2008) consider the null:

$$H_{0,MUC} : \mathbb{E}\left[I_t\left(\alpha\right)\right] = \alpha \quad \text{and} \quad \mathbb{E}\left[I_t\left(\alpha'\right)\right] = \alpha'.$$

Let us denote:

$$\begin{aligned} J_{0,t} &= 1 - J_{1,t} - J_{2,t} \\ J_{1,t} &= \begin{cases} 1 & \text{if } -VaR_{t|t-1}(\alpha') < r_t < -VaR_{t|t-1}(\alpha) \\ 0 & \text{otherwise} \end{cases} \\ J_{2,t} &= \begin{cases} 1 & \text{if } r_t < -VaR_{t|t-1}(\alpha') \\ 0 & \text{otherwise} \end{cases} . \end{aligned}$$

Testing strategies: multivariate tests (3/5)

Definition (Perignon and Smith, 2008)

The multivariate unconditional coverage test is a LR test given by:

$$\begin{aligned} LR_{MUC} &= -2\ln\left[(1-\alpha)^{H_0}\left(\alpha-\alpha'\right)^{H_1}\left(\alpha'\right)^{H_2}\right] \\ &\quad +2\ln\left[\left(1-\frac{H_0}{T}\right)^{H_0}\left(\frac{H_0}{T}-\frac{H_1}{T}\right)^{H_1}\left(\frac{H_2}{T}\right)^{H_2}\right]. \end{aligned}$$

where $H_i = \sum_{t=1}^{T} J_{i,t}$, for $i = 0, 1, 2$, denote the count variable associated with each of the Bernoulli variables.

Testing strategies: multivariate tests (3/5)

Hurlin and Tokpavi (2006):

A natural test of the CC is the univariate Ljung-Box test of

$$H_{0,CC}: r_1 = ... = r_K = 0$$

where r_k denotes the k^{th} autocorrelation:

$$LB(K) = T(T+2)\sum_{k=1}^{K}\frac{\widehat{r}_k^2}{T-k} \xrightarrow[T\to\infty]{d} \chi^2(K)$$

Testing strategies: multivariate tests (3/5)

Definition (Hurlin and Tokpavi, 2006)

Let $\Theta = \{\theta_1, .., \theta_m\}$ be a discrete set of m different coverage rates and $Hit_t = [Hit_t(\theta_1) : Hit_t(\theta_2) : ... : Hit_t(\theta_m)]'$

$$Hit_t(\theta_i) = \begin{cases} 1 - \theta_i & \text{if } r_t < -VaR_{t|t-1}(\theta_i) \\ -\theta_i & \text{else} \end{cases}$$

Under the null of CC (martingale difference):

$$H_{0,CC} : \; E\left(\left[Hit_t Hit'_{t-k}\right)\right. = 0_m \quad \forall k = 1, ..., K$$

Testing strategies

Testing strategies:

1. Frequency-based tests
2. Magnitude-based tests
3. Multivariate tests
4. Independence tests
5. Duration-based tests

Testing strategies: independence tests (4/5)

LR tests

Christoffersen (1998) assumes that the violation process $I_t(\alpha)$ can be represented as a Markov chaine with two states:

$$\Pi = \begin{pmatrix} 1-\pi_{01} & \pi_{01} \\ 1-\pi_{11} & \pi_{11} \end{pmatrix}$$

$$\pi_{ij} = \Pr\left[I_t(\alpha) = j \mid I_{t-1}(\alpha) = i \right]$$

Definition

The null of CC can be defined as follows:

$$H_{0,CC}: \ \Pi = \Pi_\alpha = \begin{pmatrix} 1-\alpha & \alpha \\ 1-\alpha & \alpha \end{pmatrix}$$

Testing strategies: independence tests (4/5)

LR tests

Christoffersen (1998) assumes that the violation process $I_t(\alpha)$ can be represented as a Markov chaine with two states:

$$\Pi = \begin{pmatrix} 1-\pi_{01} & \pi_{01} \\ 1-\pi_{11} & \pi_{11} \end{pmatrix}$$

$$\pi_{ij} = \Pr\left[I_t(\alpha) = j \mid I_{t-1}(\alpha) = i \right]$$

Definition

The null of IND can be defined as follows:

$$H_{0,IND} : \ \Pi = \Pi_{\beta} = \begin{pmatrix} 1-\beta & \beta \\ 1-\beta & \beta \end{pmatrix}$$

Testing strategies: frequency-based tests (4/5)

The corresponding LR statistics are defined by:

$$\begin{aligned} LR_{IND} &= -2\ln\left[(1-H/T)^{T-H}(H/T)^{H}\right] \\ &+2\ln\left[(1-\widehat{\pi}_{01})^{n_{00}}\widehat{\pi}_{01}^{n_{01}}(1-\widehat{\pi}_{11})^{n_{10}}\widehat{\pi}_{11}^{n_{11}}\right] \underset{T\to\infty}{\xrightarrow{d}} \chi^2(1) \end{aligned}$$

$$\begin{aligned} LR_{CC} &= -2\ln\left[(1-\alpha)^{T-H}(\alpha)^{H}\right] \\ &+2\ln\left[(1-\widehat{\pi}_{01})^{n_{00}}\widehat{\pi}_{01}^{n_{01}}(1-\widehat{\pi}_{11})^{n_{10}}\widehat{\pi}_{11}^{n_{11}}\right] \underset{T\to\infty}{\xrightarrow{d}} \chi^2(2) \end{aligned}$$

By definition:

$$LR_{CC} = LR_{UC} + LR_{IND}$$

Testing strategies: independence tests (4/5)

Table 4. Backtests of Large-Scale VaR Models

	Violation Rate	Coverage	Conditional Coverage	Independence	Serial Correlation
Bank 1	0.005	1.54 [.214]	1.57 [.455]	.0321 [.858]	-.00533 [.885]
Bank 2	0.010	.00693 [.934]	4.01 [.135]	4.00* [.046]	.158* [.016]
Bank 3	0.005	1.70 [.193]	8.81* [.012]	7.11** [.008]	.330** [.001]
Bank 4	0.000	NA	NA	NA	NA
Bank 5	0.001	8.92** [.003]	8.93* [.012]	.00271 [.959]	-.00137 [.995]
Bank 6	0.005	1.71 [.191]	1.74 [.419]	.0312 [.860]	-.00511 [.885]

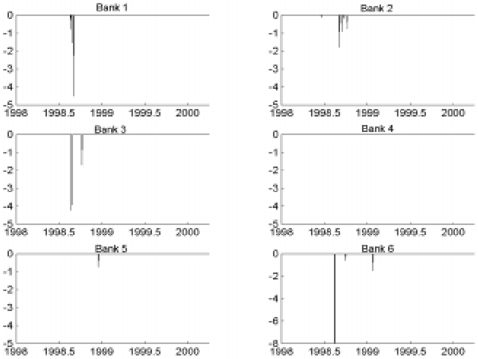
Testing strategies: independence tests (4/5)

Table 4. Backtests of Large-Scale VaR Models

	Violation Rate	Coverage	Conditional Coverage	Independence	Serial Correlation
Bank 1	0.005	1.54 [.214]	1.57 [.455]	.0321 [.858]	-.00533 [.885]
Bank 2	0.010	.00693 [.934]	4.01 [.135]	4.00* [.046]	.158* [.016]
Bank 3	0.005	1.70 [.193]	8.81* [.012]	7.11** [.008]	.330** [.001]
Bank 4	0.000	NA	NA	NA	NA
Bank 5	0.001	8.92** [.003]	8.93* [.012]	.00271 [.959]	-.00137 [.995]
Bank 6	0.005	1.71 [.191]	1.74 [.419]	.0312 [.860]	-.00511 [.885]

Testing strategies: frequency-based tests (1/5)

Figure: Violations of Banks' 99% VaR



Testing strategies: independence tests (4/5)

Regression based tests

- Engle and Manganelli (2004) suggest another approach based on a linear regression model. This model links current margin exceedances to past exceedances and/or past information.
- Let $Hit(\alpha) = I_t(\alpha) - \alpha$ be the demeaned process associated with $I_t(\alpha)$:

$$Hit_t(\alpha) = \begin{cases} 1-\alpha & \text{if } r_t < -VaR_{t|t-1}(\alpha) \\ -\alpha & \text{otherwise} \end{cases}.$$

Testing strategies: independence tests (4/5)

Regression based tests

Consider the following linear regression model:

$$Hit_t(\alpha) = \delta + \sum_{k=1}^{K} \beta_k Hit_{t-k}(\alpha) + \sum_{k=1}^{K} \gamma_k z_{t-k} + \varepsilon_t$$

where the z_{t-k} variables belong to the information set Ω_{t-1} (lagged P&L, squared past P&L, past margins, etc.)

Testing strategies: independence tests (4/5)

Regression based tests

The null hypothesis test of CC corresponds to testing the joint nullity of all the regression coefficients:

$$H_{0,CC} : \delta = \beta_k = \gamma_k = 0, \qquad \forall k = 1, ..., K.$$

since under the null :

$$\mathbb{E}\left[Hit_t\left(\alpha\right)\right] = \mathbb{E}\left[I_t\left(\alpha\right) - \alpha\right] = 0 \Longleftrightarrow \Pr\left[I_t\left(\alpha\right) = 1\right] = \alpha$$

Testing strategies: independence tests (4/5)

Definition (Engle and Manganelli, 2004)

Denote $\Psi = (\delta \ \beta_1 \ ...\beta_K \ \gamma_1 \ ...\gamma_K)'$ the vector of the $2K+1$ parameters in this model and Z the matrix of explanatory variables of model, the Wald statistic, denoted DQ_{CC}, then verifies:

$$DQ_{CC} = \frac{\widehat{\Psi}' Z' Z \widehat{\Psi}}{\alpha (1-\alpha)} \xrightarrow[T \to \infty]{d} \chi^2 (2K+1)$$

where $\widehat{\Psi}$ is the OLS estimate of Ψ.

Testing strategies: independence tests (4/5)

Regression based tests

Extension: A natural extension of the test of Engle and Manganelli (2004) consists in considering a (probit or logit) binary model linking current violations to past ones

Dumitrescu E., Hurlin C. and Pham V. (2012), Backtesting Value-at-Risk: From Dynamic Quantile to Dynamic Binary Tests, Finance

Testing strategies: independence tests (4/5)

Definition (Dumitrescu et al., 2012)

We consider a dichotomic model:

$$\Pr\left[I_t(\alpha) = 1 \mid \mathcal{F}_{t-1}\right] = F\left(\pi_t\right).$$

where $F(.)$ denotes a c.d.f. and the index π_t satisfies the following autoregressive representation:

$$\pi_t = c + \sum_{j=1}^{q_1} \beta_j \pi_{t-j} + \sum_{j=1}^{q_2} \delta_j I_{t-j}(\alpha) + \sum_{j=1}^{q_3} \gamma_j x_{t-j},$$

where $I(.)$ is a function of a finite number of lagged values of observables, and x_t is a vector of explicative variables.

Testing strategies: independence tests (4/5)

Regression based tests

$$H_0 : \beta = 0,\ \ \delta = 0, \gamma = 0 \text{ and } c = F^{-1}\left(\alpha\right).$$

since under the null of CC:

$$\Pr(I_t = 1 \mid \mathcal{F}_{t-1}) = F(F^{-1}(\alpha)) = \alpha.$$

The Dynamic Binary (DB) LR test statistic is:

$$DB_{LR_{CC}} = -2\left\{\ln L(0, F^{-1}(\alpha); I_t(\alpha), Z_t) - \ln L(\hat{\theta}, \hat{c}; I_t(\alpha), Z_t)\right\}$$
$$\xrightarrow[T\to\infty]{d} \chi^2\left(dim(Z_t)\right)$$

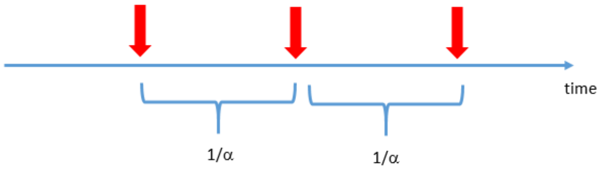
Testing strategies

Testing strategies:

1. Frequency-based tests
2. Magnitude-based tests
3. Multivariate tests
4. Independence tests
5. Duration-based tests

Testing strategies: duration-based tests (5/5)

- The UC, IND, and CC hypotheses also have some implications on the time between two consecutive VaR margin exceedances.



- Let us denote by d_v the duration between two consecutive VaR margin violations:

$$d_v = t_v - t_{v-1}$$

where t_v denotes the date of the v^{th} exceedance.

Testing strategies: duration-based tests (5/5)

Under CC hypothesis, the duration process d_v has a geometric distribution:

$$\Pr\left[d_v = k\right] = \alpha\left(1-\alpha\right)^{k-1} \qquad k \in \mathbb{N}^*.$$

- This distribution characterizes the memory-free property of the violation process $I_t(\alpha)$ with $\mathbb{E}\left(d_v\right) = 1/\alpha$

Testing strategies: duration-based tests (5/5)

Definition

Christoffersen and Pelletier (2004) use under the null hypothesis the exponential distribution:

$$g\left(d_v;\alpha\right)=\alpha\exp\left(-\alpha d_v\right).$$

Under the alternative hypothesis, they postulate a Weibull distribution for the duration variable:

$$h\left(d_v;a,b\right)=a^b b d_v^{b-1}\exp\left[-\left(a d_v\right)^b\right].$$

$$H_{0,IND}:\ b=1 \qquad H_{0,CC}:b=1,\ a=\alpha$$

Testing strategies: duration-based tests (5/5)

Drawback: we have to postulate a distribution for the duration under the alternative (misspecification of the VaR model).

Solution: Candelon et al. (2001) propose a J-test based on orthonormal polynomials associated to the geometric distribution.

Candelon B., Colletaz G., Hurlin C. et Tokpavi S. (2011), "Backtesting Value-at-Risk: a GMM duration-based test", Journal of Financial Econometrics,

Testing strategies: duration-based tests (5/5)

Candelon et al. (2001)

In the case of continuous distributions, the Pearson distributions (Normal, Student, etc.) are associated to some particular orthonormal polynomials whose expectation is equal to zero.

- These polynomials can be used as special moments to test for a distributional assumption (see. Bontemps and Meddahi, *Journal of Econometrics*, 2005).
- In the discrete case, orthonormal polynomials are defined for distributions belonging to the Ord's family (Poisson, Pascal, hypergeometric, etc.).

Testing strategies: duration-based tests (5/5)

Candelon et al. (2001)

Definition

The **orthonormal polynomials** associated to a geometric distribution with a success probability β are defined by the following recursive relationship, $\forall d \in \mathbb{N}^*$:

$$M_{j+1}(d;\beta) = \frac{(1-\beta)(2j+1)+\beta(j-d+1)}{(j+1)\sqrt{1-\beta}} M_j(d;\beta) - \left(\frac{j}{j+1}\right) M_{j-1}(d;\beta),$$

for any order $j \in \mathbb{N}$, with $M_{-1}(d;\beta) = 0$ and $M_0(d;\beta) = 1$ and:

$$\mathbb{E}\left[M_j(d;\beta)\right] = 0 \quad \forall j \in \mathbb{N}^*, \forall d \in \mathbb{N}^*.$$

Testing strategies: duration-based tests (5/5)

Candelon et al. (2001)

Example

We can show that if d follows a geometric distribution of parameter β, then:

$$M_1\left(d;\beta\right)=\left(1-\beta d\right)/\sqrt{1-\beta}$$

with

$$\mathbb{E}\left[M_1\left(d;\beta\right)\right]=0$$

Testing strategies: duration-based tests (5/5)

Candelon et al. (2001)

Our duration-based backtest procedure exploits these moment conditions.

- More precisely, let us define $\{d_1; ...; d_N\}$ a sequence of N durations between VaR violations, computed from the sequence of the hit variables $\{I_t(\alpha)\}_{t=1}^{T}$.
- Under the CC assumption, the durations d_i, $i = 1, .., N$, are $i.i.d.$ geometric(α). Hence, the null of CC can be expressed as follows:

$$H_{0,CC}: \ \mathbb{E}\left[M_j\left(d_i; \alpha\right)\right] = 0, \quad j = \{1, .., p\},$$

where p denotes the number of moment conditions.

Testing strategies: duration-based tests (5/5)

Candelon et al. (2001)

Definition

The null hypothesis of CC can be expressed as

$$H_{0,CC}: \ \mathbb{E}\left[M\left(d_i;\alpha\right)\right]=0,$$

where $M\left(d_i;\alpha\right)$ denotes a $(p,1)$ vector whose components are the orthonormal polynomials $M_j\left(d_i;\alpha\right)$, for $j=1,..,p$. Under some regularity conditions:

$$J_{CC}\left(p\right)=\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}M\left(d_i;\alpha\right)\right)^{\intercal}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}M\left(d_i;\alpha\right)\right)\xrightarrow[N\to\infty]{d}\chi^2\left(p\right)$$

Testing strategies: duration-based tests (5/5)

Candelon et al. (2001)

Definition

Under UC, the mean of durations between two violations is equal to $1/\alpha$, and the null hypothesis is

$$H_{0,UC}: \ \mathbb{E}\left[M_1\left(d_i;\alpha\right)\right]=0.$$

with a test statistic equal to

$$J_{UC}=\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}M_1\left(d_i;\alpha\right)\right)^2 \xrightarrow[N\to\infty]{d} \chi^2(1)$$

Backtesting

Recommandations

Backtesting

Recommandation 1: Test, test and test

Recommandation 2: Check the P&L data

Recommandation 3: The power of your tests may be low..

Recommandation 4: Take into account the estimation risk

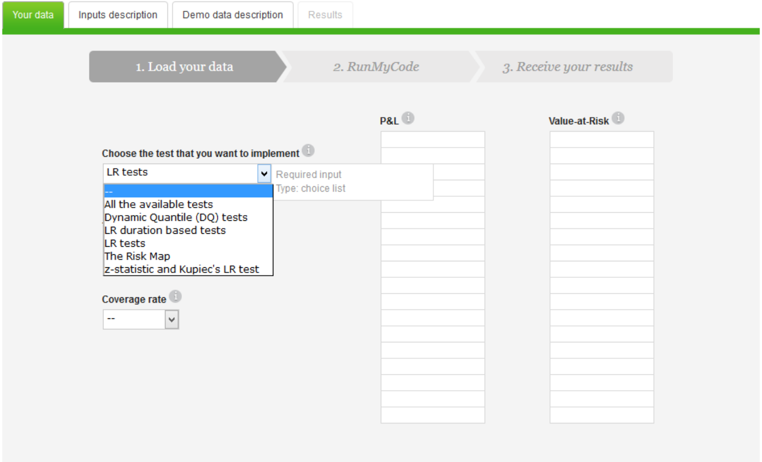
Backtesting

Recommandation 1: Test, test and test

- Each type of test (frequency, severity, independence, conditional coverage, multivariate test etc..) captures one type of potential misspecification of the VaR model.
- It is important to use a variety of tests

Backtesting

Backtesting



Backtesting

Recommandation 2: Check the P&L data

Frésard, L., C. Perignon, and W., Anders (2011), The Pernicious Effects of Contaminated Data in Risk Management, Journal of Banking and Finance.

1. A large fraction of US and international banks validate their market risk model using P&L data that include fees and commissions and intraday trading revenues.
2. Distinction between dirty P/L and hypothetical P/L (JP. Morgan, Romain Berry 2011).

Backtesting

Recommandation 3: The power of your tests may be low..

Definition

The power of a backtesting test corresponds to its capacity to detect misspecified VaR model.

$$\Pr\left[\, Rejection\ H_0 \;\middle|\; H_1 \right]$$

Example

Berkowitz, J., Christoffersen, P. F., and Pelletier, D., 2013, Evaluating Value-at-Risk Models with Desk-Level Data. *Management Science*.

Backtesting

Table 5: Power of 10% Finite Sample CC Tests on 5% VaR in Four Business Lines

Business Line 1

Sample	LB(1)	LB(5)	Markov	Weibull	Geometric	Caviar	KS	CVM
250	0.2964	0.3852	0.2048	0.1613	0.3190	0.4240	0.3485	0.3442
500	0.3912	0.5275	0.2139	0.1825	0.4466	0.5100	0.4429	0.4638
750	0.4356	0.6334	0.2257	0.2305	0.5684	0.6240	0.5316	0.5534
1000	0.4836	0.6957	0.2511	0.2698	0.6794	0.7290	0.5858	0.6068
1250	0.5431	0.7621	0.2935	0.3246	0.7560	0.7940	0.6654	0.6745
1500	0.5925	0.8146	0.3279	0.3790	0.8188	0.8590	0.7199	0.7220

Backtesting

Table 5: Power of 10% Finite Sample CC Tests on 5% VaR in Four Business Lines

Business Line 1

Sample	LB(1)	LB(5)	Markov	Weibull	Geometric	Caviar	KS	CVM
250	0.2964	0.3852	0.2048	0.1613	0.3190	0.4240	0.3485	0.3442
500	0.3912	0.5275	0.2139	0.1825	0.4466	0.5100	0.4429	0.4638
750	0.4356	0.6334	0.2257	0.2305	0.5684	0.6240	0.5316	0.5534
1000	0.4836	0.6957	0.2511	0.2698	0.6794	0.7290	0.5858	0.6068
1250	0.5431	0.7621	0.2935	0.3246	0.7560	0.7940	0.6654	0.6745
1500	0.5925	0.8146	0.3279	0.3790	0.8188	0.8590	0.7199	0.7220

Backtesting

Table 5: Power of 10% Finite Sample CC Tests on 5% VaR in Four Business Lines

Business Line 1

Sample	LB(1)	LB(5)	Markov	Weibull	Geometric	Caviar	KS	CVM
250	0.2964	0.3852	0.2048	0.1613	0.3190	0.4240	0.3485	0.3442
500	0.3912	0.5275	0.2139	0.1825	0.4466	0.5100	0.4429	0.4638
750	0.4356	0.6334	0.2257	0.2305	0.5684	0.6240	0.5316	0.5534
1000	0.4836	0.6957	0.2511	0.2698	0.6794	0.7290	0.5858	0.6068
1250	0.5431	0.7621	0.2935	0.3246	0.7560	0.7940	0.6654	0.6745
1500	0.5925	0.8146	0.3279	0.3790	0.8188	0.8590	0.7199	0.7220

Backtesting

Hurlin C. et Tokpavi S. (2008), "Une Evaluation des Procédures de Backtesting : Tout va pour le Mieux dans le Meilleur des Mondes", Finance

Idea: we use 6 different methods (GARCH, RiskMetrics, HS, CaviaR, Hybride, Delta Normale) to forecast a VaR(5%) on the same asset (GM, Nasdaq), and we apply the backtests (LR, DQ, Duration based tests) on a set of 500 samples (rolling window) of $T = 250$ forecasts.

Backtesting

Example

LR_{CC} tests: for 47% of the samples, we don't reject (at 5%) the null for any of the six VaR forecats. In 71% of the samples, we reject at the most one VaR.

Example

DQ_{CC} tests: for 20% of the samples, we don't reject (at 5%) the null for any of the six VaR forecats. In 51% of the samples, we reject at the most one VaR.

Backtesting

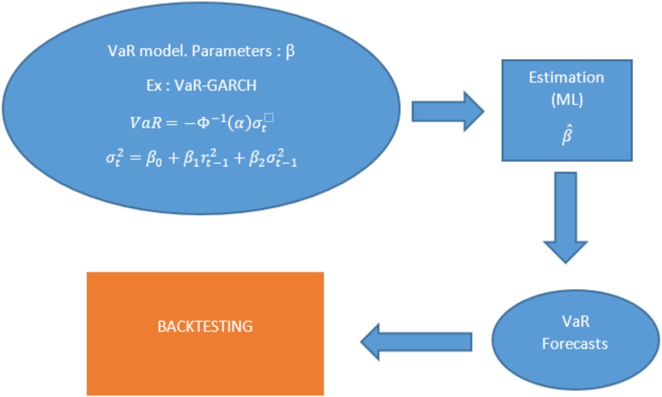
- The power of a consistent test tends to 1 when the sample size tends to inifinity.
- Recommandation: increase at the maximum the sample size of your backtest.. ($T = 500$, 750 or more.)

Backtesting

Recommandation 4: Take into account the estimation risk

- The risk dynamic is usually represented by a parametric or semi-parametric model, which has to be estimated in a preliminary step. However, the estimated counterparts of risk measures are subject to estimation uncertainty.
- Replacing, in the theoretical formulas, the true parameter value by an estimator induces a bias in the coverage probabilities.

Backtesting



Backtesting

- Escanciano and Olmo (2010, 2011) studied the effects of estimation risk on backtesting procedures. They showed how to correct the critical values in standard tests used to assess VaR models.

Escanciano, J.C. and J. Olmo (2010) Backtesting Parametric Value-at-Risk with Estimation Risk, *Journal of Business and Economics Statistics.*

Escanciano, J.C. and J. Olmo (2011) Robust Backtesting Tests for Value-at-Risk Models. *Journal of Financial Econometrics.*

Backtesting

Estimation Adjusted VaR

Gouriéroux and Zakoian (2013) a method to directly adjust the VaR to estimation risk ensuring the right conditional coverage probability at order $1/T$:

$$\Pr\left[r_t < -EVaR_{t|t-1}(\alpha)\right] = \alpha + o_P(1/T)$$

Gouriéroux C. and Zakoian J.M. (2013), Estimation Adjusted VaR, forthcoming in *Econometric Theory*.

Thank you for your attention