Quality Control of Risk Measures:
Backtesting Risk Models
“A Tale of Two Powers”*

Victor De la Peña¹, Ricardo Rivera², Jesús Ruiz-Mata³

¹ Department of Statistics, Columbia University, New York. E-mail address: vp@stat.columbia.edu
² State of New York Banking Department (NYSBD) and NYU. E-mail address: ricardo.rivera@banking.state.ny.us
³ Lehman Brothers, New York. E-mail address: jesus.ruizmata@lehman.com

*The paper on which this talk is based will appear in the January 2007 issue of The Journal of Risk.
Outline

• Quality Control problem
• VaR backtesting
• Limitations of the Basel test
• QCRM hypothesis test
• Power of the test
• New rules for accepting/rejecting VaR models
The problem

• Regulators and risk managers have to decide a course of action; i.e., accept or reject a bank’s model:

  Model correct vs. Model incorrect
VaR backtesting

• A process by which financial institutions periodically compare daily profits and losses with VaR model-generated risk measures

• The goal is to evaluate the quality and accuracy of the bank’s VaR risk model
Value at Risk: refreshment

- The $(1-\alpha) \times 100\%$ Value at Risk is the percentile $(1-\alpha)$ of the distribution of the Portfolio losses.
Exception (model failure)

- The event that the portfolio loss exceeds the corresponding VaR predicted for a trading day
Basel VaR backtest

Exceptions ($L_i > V_i$)

99% VaR model-based losses ($V_i$)

n= 250 daily observations
Notation

\( V_{i-1}^i(\alpha) \): The \((1 - \alpha) \times 100\%\) VaR estimate for trading day \( i \) using the information obtained until day \( i - 1 \),

\( L_i \): Portfolio Loss observed on day \( i \)

• The indicator of the event of an exception on day \( i \) is given by

\[
Y_i = 1 \{ L_i > V_{i-1}^i \} = \begin{cases} 
1 & \text{if } L_i > V_{i-1}^i \\
0 & \text{otherwise}
\end{cases}
\]
Assumptions

• We assume that the probabilities of observing an exception remain constant throughout time

\[ P(Y_i = 1 \mid F_{i-1}) = p, \]

where F is the information available at time t

• Technical fact: if the indicators of exceptions have the same conditional probabilities then they are independent and so

\[ X = \sum_{i=1}^{n} Y_i \approx \text{Binomial } (n, p) \]
## Basel accepting/rejecting regions

- **Green Zone** (0-4 exceptions): model is deemed accurate
- **Yellow Zone** (5-9 exceptions): Supervisor should encourage the bank to present additional information before taking action
- **Red Zone** (10+ exceptions): model is deemed inaccurate
Hypotheses

• Assume $p$ is the true (unknown) probability of having an exception, risk managers test

\[ H_0: p = p_0 = 0.01 \quad \text{vs.} \quad H_A: p > p_0 = 0.01 \]

• where $p_0 = 0.01$ (99% VaR) is the probability of an exception when the model is correct
Control Type I Error

• Basel VaR backtesting method seeks to control the probability of rejecting the VaR model when it is correct

  • Set the probability of rejecting the VaR model when it is correct to be as small as 0.0003 (0.03%)

  • Therefore, it controls the type I error at 0.03%

  • \( P(\text{number of exceptions} \geq 10 \text{ when } p = 0.01) = 0.0003 \)
Basel on VaR Backtesting

“The Committee of course recognizes that tests of this type are limited in their power to distinguish an accurate model from an inaccurate model”\(^1\)

(1) Basel Committee on Banking Supervision (Basel), page 5 of “Supervisory Framework for the use of “Back Testing” in conjunction with the internal models approach to Market Risk Capital requirements”, January 1996
Change of hypotheses

• QCRM hypothesis testing problem:

H₀: VaR Model incorrect vs. Hₐ: VaR Model correct

• Accepting H₀ implies rejecting the model
• Rejecting H₀ implies accepting the model
New hypothesis test

• Assume $p$ is the true probability of having one exception (unknown), QCRM tests:

$$H_0^Q: p > 0.01 \quad \text{vs.} \quad H_A^Q: p \leq 0.01$$

• This is the quality control problem
New acceptance and rejection regions

- **New Green zone = {0 to 5 exceptions}:** if $p_0$ is in the 95% one-sided confidence interval for $p [p_L(x,.05), 1]$

- **New Yellow zone = {6 or 7 exceptions}:** if $p_0$ is in the 99% one-sided confidence interval for $p [p_L(x,.01), 1]$ (and it is not in the 95% one-sided confidence interval)

- **New Red Zone = {8 or more exceptions}:** if $p_0$ is not in the 99% one-sided confidence interval for $p [p_L(x,.01), 1]$
Look at the power of the test!

• The power of the test is a function of the (unknown) parameter $p$, which is defined in terms of the rejection region $R$ as

$$\beta(p) = P_p(X \in R)$$

• This function contains all the information about the QCRM test

• We redefine the power of the test in terms of probability of accepting (rejecting) an incorrect (correct) model
Power: key comparison

| Tests   | P(rejecting the model|correct) | P(rejecting the model|incorrect)  |
|---------|----------------------|------------------|
| Basel   | 0 – 0.0003*          | P(X≥10|given p>0.01) |
| QCRM    | 0 – 0.004            | P(X≥8|given p>0.01) |

* Assume composite null hypothesis for Basel test with p≤0.01
Idea

• QCRM increases, with respect to the Basel test, the probability of rejecting an incorrect model

• QCRM’s null hypothesis is then rejected when there is overwhelming evidence to accept the model ⇒

• This lead to an statistically certification of the model
Probability of rejecting a correct model

- Basel: [0 – 0.0003] and QCRM [0 – 0.004]

- Suppose 10 model reviews per year. How many years, on average, are necessary for regulators to make a wrong assessment?…

<table>
<thead>
<tr>
<th>Test</th>
<th>Max. Error</th>
<th>Model Reviews</th>
<th>Years</th>
<th>Years per Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basel</td>
<td>3</td>
<td>10,000</td>
<td>1,000</td>
<td>333.3</td>
</tr>
<tr>
<td>QCRM</td>
<td>4</td>
<td>1,000</td>
<td>100</td>
<td>25</td>
</tr>
</tbody>
</table>
Probability of rejecting a wrong model

X-axis: different values of alternative hypotheses $p$

$P(X \geq 8 \mid p = p_0 > 0.01)$

$P(X \geq 10 \mid p = p_1 > 0.01)$

QCRM

Basel
Power rate curve

- Percentage gains of QCRM over Basel in the probability of rejecting the wrong model for different values of the alternative hypotheses $p$
Research in progress

- QCRM to test credit risk models for Basel II implementation
- The test can be applied to other areas within or outside finance
Summary

• We find that the Basel test is extremely conservative; i.e., it almost guarantees that regulators will not reject a correct model

• …but it may lead regulators to accept an incorrect model

• We propose a more balanced test that dramatically increases, with respect to Basel, the probability of rejecting a wrong model

• We propose new rules for accepting/rejecting a VaR model

• We can use QCRM to test the validity of credit risk models for Basel II implementation
References

- Supervisory Framework for the Use of “Back testing” in conjunction with the internal models approach to Market Risk Capital requirements. Basle Committee on Banking Supervision
Preambulo: Riesgo de Mercado

• Que es el riesgo de mercado?
• Acuerdo de Basel
• Herramientas usadas
  • Binomial
  • Modelos de VaR (Value-at-Risk)
  • Teoria de pruebas de hipotesis
Que es el riesgo de mercado?

- Riesgo de perdidas en el portafolio del banco debido a cambios en los precios de los activos financieros

- Portafolio: conjunto de inversiones del banco en activos financieros

- Activos financieros incluye: acciones, bonos, prestamos, derivados, etc.

- Riesgo de credito: es el riesgo potencial de perdidas debido a la bancarrota de los deudores del banco
Acuerdo de Basel

• Basel es un organismo internacional dedicado a establecer normas para la “mejor practica” del manejo y control de los riesgos bancarios

• Basel establecio las normas para el uso de modelos internos (matematicos) de los bancos para la medicion y administracion del riesgo de mercado

• En 1996 establecio las reglas para la validacion de los modelos internos de los bancos, las que son utilizadas a nivel internacional
Herramientas usadas

- Binomial

\[ P(X = k) = \frac{n!}{k!(n-k)!} p^k (1 - p)^{n-k} \]

- Ejemplo: cual es la probabilidad de obtener “cara” 4 veces al tirar una moneda 10 veces?

\[ P(X = 4) = \frac{10!}{4!(10-4)!} 0.5^4 (1 - 0.5)^{10-4} = 0.205078 \]
Aplicacion: Valor a Riesgo (VaR)

• El $(1-\alpha)\times100\%$ VaR es el quantil $(1-\alpha)$ de la distribución
  de las perdidas del portafolio del banco
VaR backtesting

• Es el proceso por el cual los bancos comparan periodicamente sus perdidas y ganancias diarias con los valores generados mediante el uso del modelo VaR

• El objetivo es el evaluar la calidad de las predicciones del modelo VaR
Perdidas ($)

32

Ganancias ($)

0

99% modelo de VaR ($V_i$)

excepciones ($P_i > V_i$)

n= 250 observaciones diarias

Basel VaR backtest
Hipótesis de la prueba de Basel

- Supongamos que $p$ es la verdadera probabilidad de cometer un error (excepción)

\[ H_0: p = p_0 = 0.01 \quad \text{vs.} \quad H_A: p > p_0 = 0.01 \]

- Donde $p_0 = 0.01$ (99% VaR) es la probabilidad de cometer un error cuando el modelo es correcto.

- $n$ es igual a 250 observaciones.

- $k$ el número de excepciones es mayor o igual a 10.

- $P(X \geq 10) = 0.0003$