The validation of Credit Rating and Scoring Models

Raffaella Calabrese
raffaella.calabrese1@unimib.it
University of Milano-Bicocca

Swiss Statistics Meeting
Geneva, Switzerland
October 29th, 2009
Outline

1. The validation process
2. Literature review
   - Cumulative Accuracy Profile Curve
   - Receiver Operating Characteristic Curve
3. Methodological proposals
   - Curve of Classification Error Costs and Error Costs
1 The validation process

2 Literature review
   - Cumulative Accuracy Profile Curve
   - Receiver Operating Characteristic Curve

3 Methodological proposals
   - Curve of Classification Error Costs and Error Costs
Outline

1. The validation process

2. Literature review
   - Cumulative Accuracy Profile Curve
   - Receiver Operating Characteristic Curve

3. Methodological proposals
   - Curve of Classification Error Costs and Error Costs
Credit **rating** and **scoring** models estimate the credit obligor’s **worthiness** and provide an assessment of the obligor’s future status.

The **discriminatory power** of a rating or scoring model denotes its ability to discriminate **ex ante** between defaulting and non-defaulting borrowers.

The **validation** process assesses the discriminatory power of a rating or scoring model.
Credit **rating** and **scoring** models estimate the credit obligor’s **worthiness** and provide an assessment of the obligor’s future status.

The **discriminatory power** of a rating or scoring model denotes its ability to discriminate *ex ante* between defaulting and non-defaulting borrowers.

The **validation** process assesses the discriminatory power of a rating or scoring model.
Credit **rating** and **scoring** models estimate the credit obligor’s *worthiness* and provide an assessment of the obligor’s future status.

The **discriminatory power** of a rating or scoring model denotes its ability to discriminate *ex ante* between defaulting and non-defaulting borrowers.

The **validation** process assesses the discriminatory power of a rating or scoring model.
Each borrower is characterized by two random variables:

- the score $S$ assigned to the borrower is a continuous r. v. with support $(-\infty, \infty)$
- the Bernoulli r.v. $B$ represents the borrower’s state at the end of a fixed time-period

\[
B = \begin{cases} 
1, & \text{the borrower’s state is default (d)}; \\
0, & \text{the borrower’s state is non default (n)}. 
\end{cases}
\]

The conditional distribution functions of $S$ given a value of $B$ are denoted respectively by $F_d(\cdot)$ and $F_n(\cdot)$.
Each borrower is characterized by two random variables:

- The score $S$ assigned to the borrower is a continuous r. v. with support $(-\infty, \infty)$
- The Bernoulli r.v. $B$ represents the borrower’s state at the end of a fixed time-period

$$B = \begin{cases} 
1, & \text{the borrower’s state is default (d)}; \\
0, & \text{the borrower’s state is non default (n)}. 
\end{cases}$$

The conditional distribution functions of $S$ given a value of $B$ are denoted respectively by $F_d(\cdot)$ and $F_n(\cdot)$. 
Each borrower is characterized by two random variables:

- the score $S$ assigned to the borrower is a continuous r. v. with support $(-\infty, \infty)$
- the Bernoulli r.v. $B$ represents the borrower’s state at the end of a fixed time-period

$$B = \begin{cases} 1, & \text{the borrower’s state is default (d)}; \\ 0, & \text{the borrower’s state is non default (n)}. \end{cases}$$

The conditional distribution functions of $S$ given a value of $B$ are denoted respectively by $F_d(\cdot)$ and $F_n(\cdot)$. 
The distribution function of the score $S$ is

$$F(s) = pF_d(s) + (1 - p)F_n(s)$$

where $p$ is the probability of default $p = P[B = d]$.

The accuracy (AC) is

$$AC = pF_d(s) + (1 - p)[1 - F_n(s)] = 2pF_d(s) - F(s) + (1 - p)$$
The distribution function of the score $S$ is

$$F(s) = pF_d(s) + (1 - p)F_n(s)$$

where $p$ is the probability of default $p = P[B = d]$. The accuracy (AC) is

$$AC = pF_d(s) + (1 - p)[1 - F_n(s)] = 2pF_d(s) - F(s) + (1 - p)$$
<table>
<thead>
<tr>
<th></th>
<th>Actual default</th>
<th>Actual non default</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Predicted default</strong> (below s)</td>
<td>True Default (TD)</td>
<td>False Default (FD)</td>
</tr>
<tr>
<td><strong>Predicted non default</strong> (above s)</td>
<td>False Non Default (FN)</td>
<td>True Non Default (TN)</td>
</tr>
</tbody>
</table>

- hit rate \( \hat{F}_d(s) = \frac{TD}{N_d} \)
- false alarm rate \( \hat{F}_n(s) = \frac{FD}{N_n} \)
- \( \hat{F}_d(s) = \hat{p} \hat{F}_d(s) + (1 - \hat{p}) \hat{F}_n(s) = \frac{TD + FD}{N_d + N_n} \)

where \( \hat{p} = \frac{N_d}{N_d + N_n} \)
<table>
<thead>
<tr>
<th><strong>Predicted default</strong> (below s)</th>
<th><strong>Actual default</strong></th>
<th><strong>Actual non default</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>True Default (TD)</td>
<td></td>
<td>False Default (FD)</td>
</tr>
<tr>
<td>Type II error</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Predicted non default</strong> (above s)</td>
<td>False Non Default (FN)</td>
<td>True Non Default (TN)</td>
</tr>
<tr>
<td>Type I error</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th><strong>True Default (TD)</strong></th>
<th><strong>False Default (FD)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N_d$</td>
<td>$N_n$</td>
</tr>
</tbody>
</table>

- Hit rate $\hat{F}_d(s) = \frac{TD}{N_d}$
- False alarm rate $\hat{F}_n(s) = \frac{FD}{N_n}$

$\hat{F}(s) = \hat{p}\hat{F}_d(s) + (1 - \hat{p})\hat{F}_n(s) = \frac{TD + FD}{N_d + N_n}$

where $\hat{p} = \frac{N_d}{N_d + N_n}$
<table>
<thead>
<tr>
<th>Predicted default (below s)</th>
<th>Actual default</th>
<th>Actual non default</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Default (TD)</td>
<td>False Default (FD)</td>
<td></td>
</tr>
<tr>
<td><em>Type II error</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>False Non Default (FN)</td>
<td>True Non Default (TN)</td>
<td></td>
</tr>
<tr>
<td><em>Type I error</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Predicted non default (above s)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>False Default (FD)</td>
<td></td>
</tr>
<tr>
<td><em>Type II error</em></td>
<td></td>
</tr>
</tbody>
</table>

- Hit rate: \( \hat{F}_d(s) = \frac{TD}{Nd} \)
- False alarm rate: \( \hat{F}_n(s) = \frac{FD}{Nn} \)

\[
\hat{F}(s) = \hat{p} \hat{F}_d(s) + (1 - \hat{p}) \hat{F}_n(s) = \frac{TD + FD}{Nd + Nn}
\]

where \( \hat{p} = \frac{Nd}{Nd + Nn} \)
### The validation process

#### Literature review

#### Methodological proposals

#### Cumulative Accuracy Profile Curve

#### Receiver Operating Characteristic Curve

<table>
<thead>
<tr>
<th>Predicted default (below s)</th>
<th>Actual default</th>
<th>Actual non default</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Default (TD)</td>
<td></td>
<td>False Default (FD)</td>
</tr>
<tr>
<td>Type II error</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Predicted non default (above s)</th>
<th>False Non Default (FN)</th>
<th>True Non Default (TN)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I error</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- $\hat{F}_d(s) = \frac{TD}{N_d}$
- $\hat{F}_n(s) = \frac{FD}{N_n}$
- $\hat{F}(s) = \hat{p}\hat{F}_d(s) + (1 - \hat{p})\hat{F}_n(s) = \frac{TD + FD}{N_d + N_n}$
- $\hat{p} = \frac{N_d}{N_d + N_n}$

Where

- $N_d$: Number of default observations
- $N_n$: Number of non-default observations
Cumulative Accuracy Profile (CAP) Curve and Accuracy Ratio (AR)

curve: $\text{CAP}(u) = F_d[F^{-1}(u)], \quad u \in (0, 1)$
Cumulative Accuracy Profile (CAP) Curve and Accuracy Ratio (AR)

\[
\text{CAP}(u) = F_d[F^{-1}(u)], \quad u \in (0, 1)
\]
Cumulative Accuracy Profile (CAP) Curve and Accuracy Ratio (AR)

- **synthetic index** (BCBS, 2005):
  \[ AR = \frac{a_R}{a_R + a_Q} \quad AR \in [0, 1] \]

- **optimal cut-off score** (Hong, 2009): the intersection of the CAP curve and the iso-performance tangent line
  \[ F_d(s) = \frac{1}{2p} [F(s) + AC + p - 1] \]

- **drawbacks**:
  - dependence on the sample relative frequency of defaulted borrowers;
  - the type II error and the costs of wrong classification are ignored.
Cumulative Accuracy Profile (CAP) Curve and Accuracy Ratio (AR)

- **synthetic index** (BCBS, 2005):
  \[ AR = \frac{a_R}{a_R + a_Q} \quad AR \in [0, 1] \]

- **optimal cut-off score** (Hong, 2009): the intersection of the CAP curve and the iso-performance tangent line
  \[ F_d(s) = \frac{1}{2p} [F(s) + AC + p - 1] \]

**drawbacks:**
- dependence on the sample relative frequency of defaulted borrowers;
- the type II error and the costs of wrong classification are ignored.
Cumulative Accuracy Profile (CAP) Curve and Accuracy Ratio (AR)

- **synthetic index** (BCBS, 2005):
  \[ AR = \frac{a_R}{a_R + a_Q} \quad AR \in [0, 1] \]

- **optimal cut-off score** (Hong, 2009): the intersection of the CAP curve and the iso-performance tangent line
  \[ F_d(s) = \frac{1}{2p} \left[ F(s) + AC + p - 1 \right] \]

- **drawbacks:**
  - dependence on the sample relative frequency of defaulted borrowers;
  - the type II error and the costs of wrong classification are ignored.
Cumulative Accuracy Profile (CAP) Curve and Accuracy Ratio (AR)

- **synthetic index** (BCBS, 2005):
  \[ AR = \frac{a_R}{a_R + a_Q} \quad AR \in [0, 1] \]

- **optimal cut-off score** (Hong, 2009): the intersection of the CAP curve and the iso-performance tangent line
  \[ F_d(s) = \frac{1}{2p} [F(s) + AC + p - 1] \]

- **drawbacks:**
  - dependence on the sample relative frequency of defaulted borrowers;
  - the type II error and the costs of wrong classification are ignored.
Receiver Operating Characteristic (ROC) Curve and Area Under the Curve (AUC)

Curve: \( ROC(u) = F_d[F^{-1}(u)], \quad u \in (0, 1) \)
Receiver Operating Characteristic (ROC) Curve and Area Under the Curve (AUC)

Curve: \( ROC(u) = F_d[F^{-1}(u)] \), \( u \in (0, 1) \)
**Synthetic index** (BCBS, 2005):

\[
AUC = \int_0^1 F_d[F_n(s)]dF_n(s) \quad \text{AUC} \in [0.5, 1]
\]

**Optimal cut-off score** (Hong, 2009): the intersection of the ROC curve and the iso-performance tangent line

\[
F_d(s) = \frac{1 - p}{p}F_n(s) + \frac{1}{p}(AC + p - 1).
\]

**Drawbacks**: the costs of wrong classification are ignored.
Synthetic index (BCBS, 2005):

\[ AUC = \int_{0}^{1} F_d[F_n(s)]dF_n(s) \quad AUC \in [0.5, 1] \]

Optimal cut-off score (Hong, 2009): the intersection of the ROC curve and the iso-performance tangent line

\[ F_d(s) = \frac{1 - p}{p}F_n(s) + \frac{1}{p}(AC + p - 1). \]

drawbacks: the costs of wrong classification are ignored.
Synthetic index (BCBS, 2005):

\[
AUC = \int_{0}^{1} F_d[F_n(s)]dF_n(s) \quad AUC \in [0.5, 1]
\]

Optimal cut-off score (Hong, 2009): the intersection of the ROC curve and the iso-performance tangent line

\[
F_d(s) = \frac{1 - p}{p} F_n(s) + \frac{1}{p} (AC + p - 1).
\]

drawbacks: the costs of wrong classification are ignored.
Curve of Classification Error Costs (CEC) and Error Costs (EC)

**Curve:**

\[
C[u] = \frac{C_{FN}}{2} \left\{ 1 - F_d[F^{-1}(u)] \right\} + \frac{C_{FD}}{2} F_n[F^{-1}(u)] \quad u \in (0, 1)
\]

**Synthetic index:**

\[
EC^* = \int_0^1 C[F(s)] dF(s)
\]

\[
EC = \frac{EC^* - EC_R^*}{EC_P^* - EC_R^*} \quad EC \in [0, 1]
\]

where \( EC_R^* \) and \( EC_P^* \) are respectively the error costs of the random and perfect models.
Curve of Classification Error Costs (CEC) and Error Costs (EC)

- **Curve:**
  \[
  C[u] = \frac{C_{FN}}{2} \{1 - F_d[F^{-1}(u)]\} + \frac{C_{FD}}{2} F_n[F^{-1}(u)] \quad u \in (0, 1)
  \]

- **Synthetic index:**
  \[
  EC^* = \int_0^1 C[F(s)] dF(s)
  \]
  \[
  EC = \frac{EC^* - EC^*_R}{EC^*_P - EC^*_R} \quad EC \in [0, 1]
  \]

where \( EC^*_R \) and \( EC^*_P \) are respectively the error costs of the random and perfect models.
Curve of Classification Error Costs (CEC) and Error Costs (EC)

**Optimal cut-off score:** the value $s$ that satisfies

$$
\min_s \left\{ \frac{C_{FN}}{2} [1 - F_d(s)] + \frac{C_{FD}}{2} F_n(s) \right\} = \max_s \left[ \frac{F_d(s)}{C_{FD}} - \frac{F_n(s)}{C_{FN}} \right]
$$

**Point measure** (Zenga, 2007): $U(c) = \frac{\bar{\mu}(c)}{\mu^\dagger(c)}$

where $c \in (-\infty, +\infty)$ and

$$
\bar{\mu}(c) = \frac{1}{F(c)} \int_{-\infty}^{c} \left\{ \frac{C_{FN}}{2} [1 - F_d(s)] + \frac{C_{FD}}{2} F_n(s) \right\} dF(s)
$$

$$
\mu^\dagger(c) = \frac{1}{1 - F(c)} \int_{c}^{+\infty} \left\{ \frac{C_{FN}}{2} [1 - F_d(s)] + \frac{C_{FD}}{2} F_n(s) \right\} dF(s)
$$

Raffaella Calabrese
Validation of internal rating systems
**Curve of Classification Error Costs (CEC) and Error Costs (EC)**

- **Optimal cut-off score**: the value $s$ that satisfies

  $$
  \min_s \left\{ \frac{C_{FN}}{2} [1 - F_d(s)] + \frac{C_{FD}}{2} F_n(s) \right\} = \max_s \left[ \frac{F_d(s)}{C_{FD}} - \frac{F_n(s)}{C_{FN}} \right]
  $$

- **Point measure** (Zenga, 2007): $U(c) = \frac{\bar{\mu}(c)}{\hat{\mu}(c)}$

  where $c \in (-\infty, +\infty)$ and

  $$
  \bar{\mu}(c) = \frac{1}{F(c)} \int_{-\infty}^{c} \left\{ \frac{C_{FN}}{2} [1 - F_d(s)] + \frac{C_{FD}}{2} F_n(s) \right\} dF(s)
  $$

  $$
  \hat{\mu}(c) = \frac{1}{1 - F(c)} \int_{c}^{+\infty} \left\{ \frac{C_{FN}}{2} [1 - F_d(s)] + \frac{C_{FD}}{2} F_n(s) \right\} dF(s)
  $$
