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

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

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

Preliminary definitions

- 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

Preliminary definitions

- 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

Preliminary definitions

- 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

$$AC = pF_d(s) + (1-p)[1-F_n(s)] = 2pF_d(s) - F(s) + (1-p)$$

	Actual default	**Actual non default**
Predicted default (below s)	True Default (TD)	False Default (FD) *Type II error*
Predicted non default (above s)	False Non Default (FN) *Type I error*	True Non Default (TN)
	N_d	N_n

- 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 $\hat{p} = \frac{N_d}{N_d+N_n}$

	Actual default	**Actual non default**
Predicted default (below s)	True Default (TD)	False Default (FD) *Type II error*
Predicted non default (above s)	False Non Default (FN) *Type I error*	True Non Default (TN)
	N_d	N_n

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

  where $\hat{p} = \dfrac{N_d}{N_d + N_n}$

	Actual default	**Actual non default**
Predicted default (below s)	True Default (TD)	False Default (FD) *Type II error*
Predicted non default (above s)	False Non Default (FN) *Type I error*	True Non Default (TN)
	N_d	N_n

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

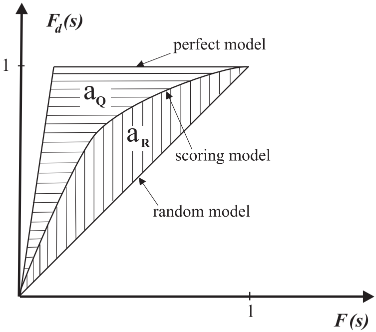
  where $\hat{p} = \frac{N_d}{N_d + N_n}$

	Actual default	**Actual non default**
Predicted default (below s)	True Default (TD)	False Default (FD) *Type II error*
Predicted non default (above s)	False Non Default (FN) *Type I error*	True Non Default (TN)
	N_d	N_n

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

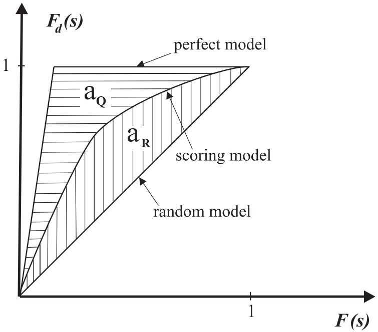
  where $\hat{p} = \dfrac{N_d}{N_d+N_n}$

Cumulative Accuracy Profile (CAP) Curve and Accuracy Ratio (AR)



- curve: $CAP(u) = F_d[F^{-1}(u)], \qquad u \in (0,1)$

Cumulative Accuracy Profile (CAP) Curve and Accuracy Ratio (AR)



- *curve*: $CAP(u) = F_d[F^{-1}(u)], \qquad u \in (0,1)$

Cumulative Accuracy Profile (CAP) Curve and Accuracy Ratio (AR)

- *synthetic index* (BCBS, 2005):

$$AR = \frac{a_R}{a_R + a_Q} \qquad AR \in [0, 1]$$

- *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):

$$AR = \frac{a_R}{a_R + a_Q} \qquad AR \in [0, 1]$$

- *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):

$$AR = \frac{a_R}{a_R + a_Q} \qquad AR \in [0, 1]$$

- *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):

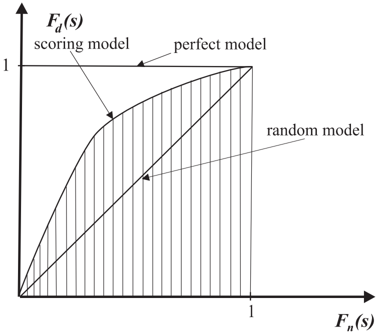
$$AR = \frac{a_R}{a_R + a_Q} \qquad AR \in [0, 1]$$

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

$$F_d(s) = \frac{1}{2p}\left[F(s) + AC + p - 1\right]$$

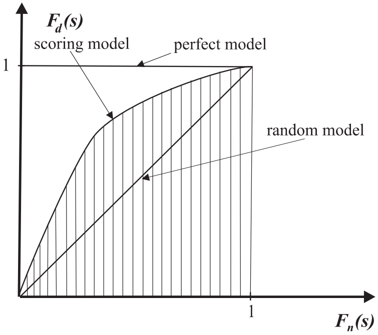
- *drawbacks*:
  - dependence on the sample relative frequency of defaulted borrowers;
  - the type II error and the costs of wrong classification are ignored.

Receiver Operating Characteristic (ROC) Curve and Area Under the Curve (AUC)



- Curve: $ROC(u) = F_d[F^{-1}(u)], \qquad u \in (0,1)$

Receiver Operating Characteristic (ROC) Curve and Area Under the Curve (AUC)



- *Curve*: $ROC(u) = F_d[F^{-1}(u)], \qquad u \in (0,1)$

- *Synthetic index* (BCBS, 2005):

$$AUC = \int_0^1 F_d[F_n(s)]dF_n(s) \qquad AUC \in [0.5, 1]$$

- *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):

$$AUC = \int_0^1 F_d[F_n(s)]dF_n(s) \qquad AUC \in [0.5, 1]$$

- *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):

$$AUC = \int_0^1 F_d[F_n(s)]dF_n(s) \qquad AUC \in [0.5, 1]$$

- *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*:

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

- *Synthetic index*:

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

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

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*:

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

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

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): $U(c) = \frac{\bar{\mu}(c)}{\overset{+}{\mu}(c)}$

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

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

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

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): $U(c) = \frac{\bar{\mu}(c)}{\overset{+}{\mu}(c)}$

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

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

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

Bibliography

- Basel Committee on Banking Supervision (2005). *Studies on the Validation of Internal Rating Systems*. Working paper 14. Basel, BIS.
- Hong C. S. (2009). *Optimal Threshold from ROC and CAP Curves.* Communications in Statistics, 38, 2060-2072.
- Zenga M. (2007). *Inequality curve and inequality index based on the ratio between lower and upper arithmetic means.* Statistica & Applicazioni, V, 3-27.