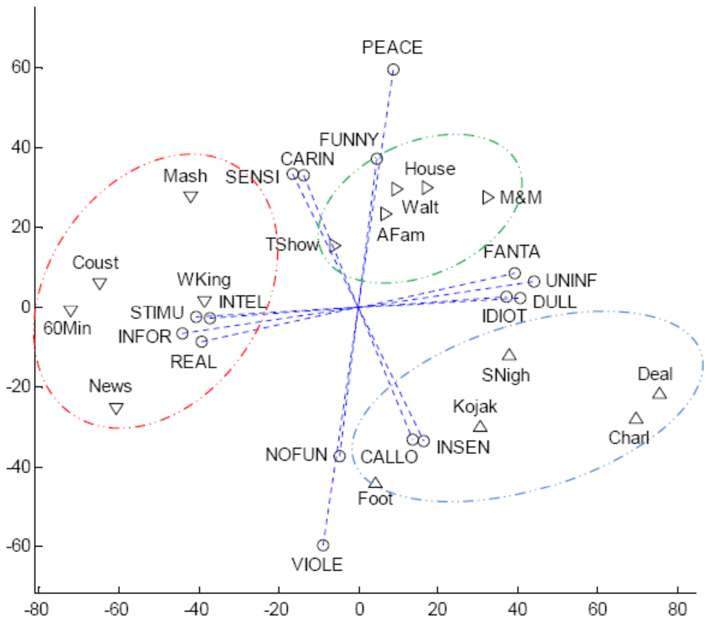


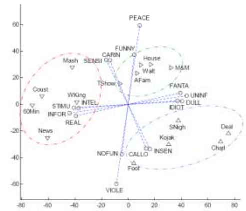
Metodi Statistici per il Management

Applicazioni: Mercati Finanziari e Banking

Simone Borra - Roberto Rocci

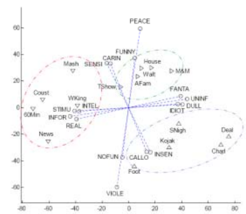
Indice

– Tipi di rischio nelle banche
– Basilea II e il requisito patrimoniale
– Obiettivi della gestione dei rischi
– Rischio di mercato
– Rischio di credito
– Rischio operativo



Tipi di rischio

- ***Rischio di mercato***. È il rischio di perdite sul *portafoglio di proprietà* della banca.
- ***Rischio di credito***. È il rischio di perdite dovuto ai mancati pagamenti da parte del debitore, o da ritardi negli stessi.
- ***Rischio operativo***. È il rischio di “perdite dovute all’inadeguatezza o fallimento di processi, persone, sistemi”.

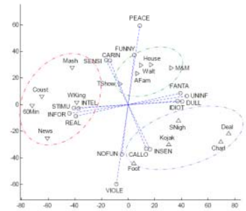


Obiettivi della normativa

•***Individuazione e classificazione dei rischi***.

•***Metodi di misura del rischio***. Il comitato ha tutto l'interesse a incentivare metodi di misura del rischio di tipo statistico, in quanto dovrebbero portare a una maggiore precisione e consapevolezza.

•***Gestione del rischio***.

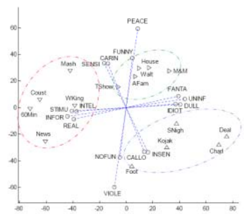


Basilea II: il requisito patrimoniale

Il patrimonio della banca, detto *Patrimonio di Vigilanza*, deve essere superiore ad una soglia minima, che dipende appunto dall'ammontare dei rischi. Attualmente:

$$\textit{\textbf{Patrimonio di vigilanza / Misura Rischi}} \geq \textit{\textbf{8\%}}$$

La misura totale dei rischi è detta più comunemente *capitale assorbito* o *requisito patrimoniale*.

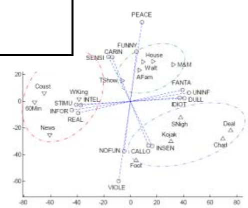


Tipi di rischio e previsione normativa

Rischio \ Normativa	Basilea II	Basilea I
Rischio di mercato	▪ Metodo standard ▪ Modelli interni	▪ Metodo standard ▪ Modelli interni

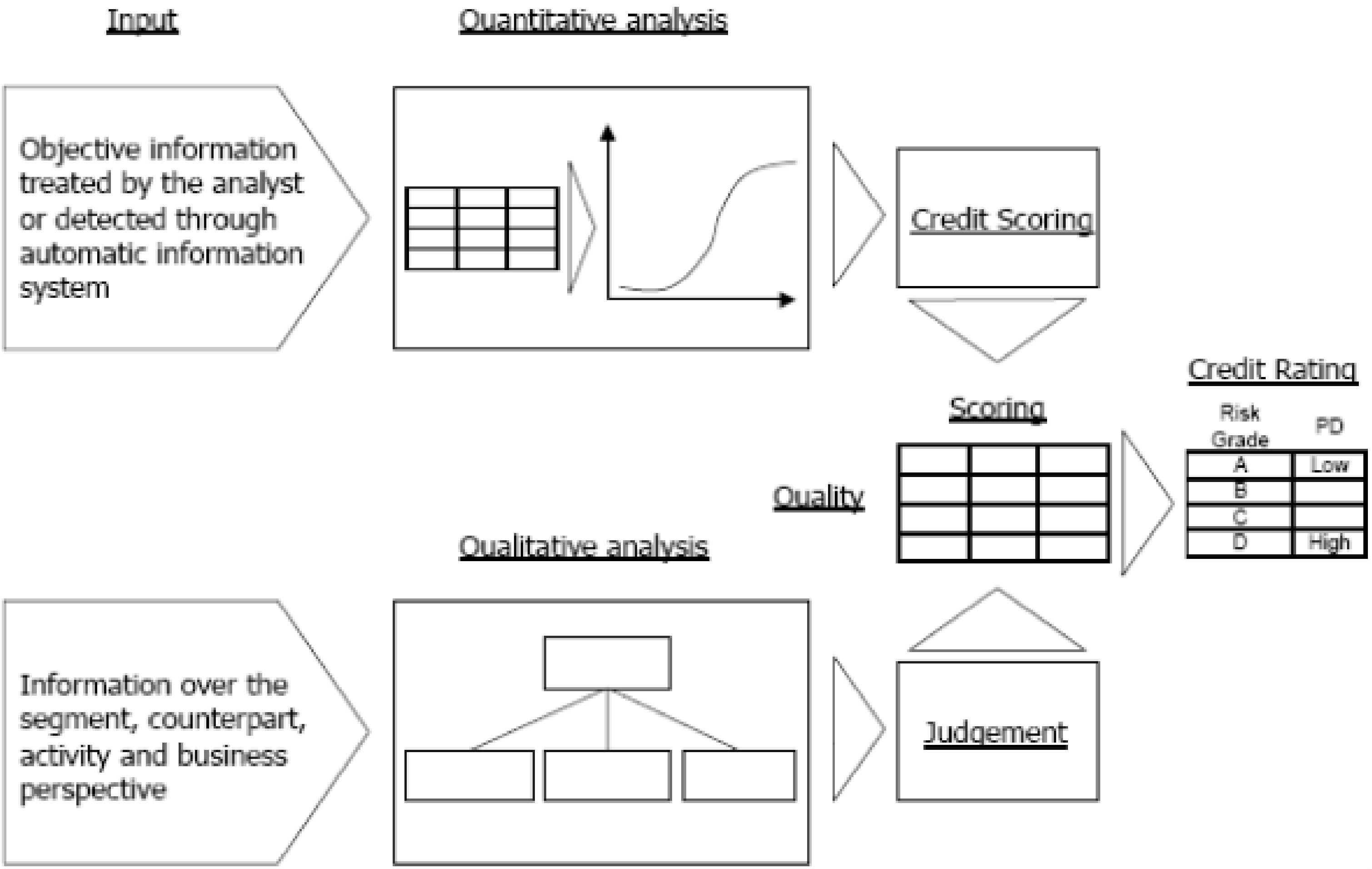
Rischio di credito	▪ Metodo standard
	▪ Modelli interni (IRB)
	  ○ Foundation
	  ○ Advanced
		▪ Metodo standard

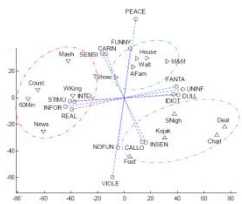
Rischio operativo	▪ Metodo *basic*
	▪ Metodo standard
	▪ Modello interno AMA
	  ○ Gamma
	  ○ LDA
		Rischio non considerato



Modelli di Scoring

Sistema di RATING = processo che consente di elaborare una valutazione sintetica (su scala ordinale) della rischiosità di un cliente





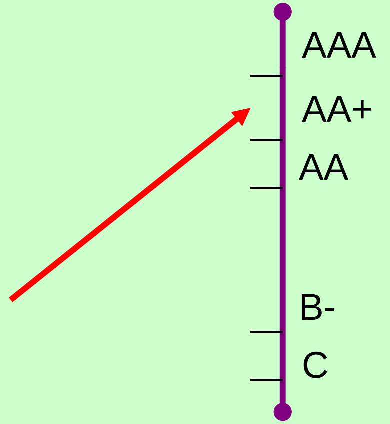
Modelli di Scoring

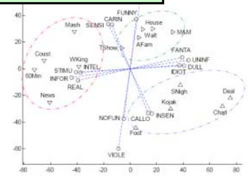
Nella loro review sulle metodologie per la misurazione del rischio di credito, Altman e Saunders (1998) considerano quattro metodologie:

- modello lineare di probabilità;
- modello logit;
- modello probit;
- analisi discriminante.

Ci occuperemo del modello logit, oggi considerato "best practice"

Indicatori:	I_1	I_2	…	I_p
Pesi:	a_1	a_2	…	a_p
Scoring:	$S = I_1 a_1 + I_2 a_2 +$		…	$+ I_p\, a_p$





Esempio: variabili

Una banca vuole capire la relazione che intercorre tra

Y fattore a 2 livelli: *buen – mal*, se restituisce il prestito rispettando i termini;

e

Cuenta fattore 3 levelli: *no - good running - bad running*, qualità del conto corrente;

Mes durata del prestito in mesi;

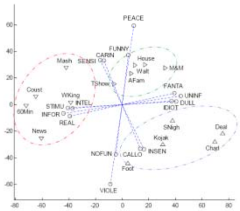
Ppag fattore 2 levelli: *pre buen pagador - pre mal pagador*, storia passata del cliente;

Uso fattore 2 livelli *privado - profesional*, utilizzo del prestito;

DM ammontare del prestito in marchi;

Sexo fattore 2 levelli: *mujer - hombre*, sesso;

Estc fattore 2 levelli: *no vive solo - vive solo*, stato civile.



Esempio: dati

y	Frequency	Percent
buen	700	70
mal	300	30

cuenta	Frequency	Percent
bad running	332	33.2
good running	394	39.4
no	274	27.4

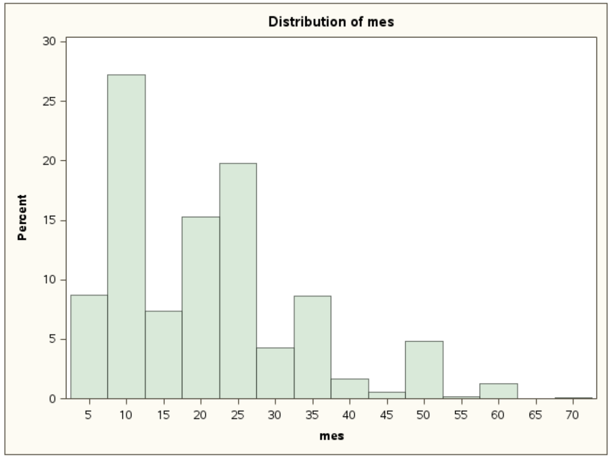
sexo	Frequency	Percent
hombre	598	59.8
mujer	402	40.2

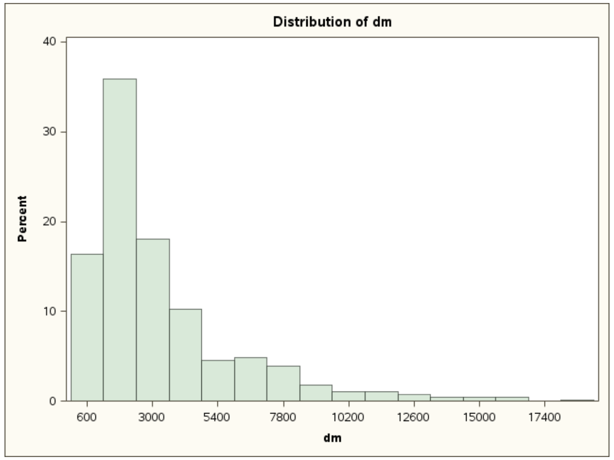
uso	Frequency	Percent
privado	657	65.7
profesional	343	34.3

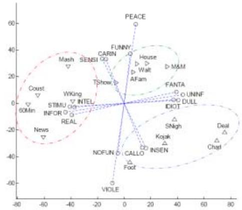
ppag	Frequency	Percent
pre buen paga	911	91.1
pre mal paga	89	8.9

estc	Frequency	Percent
no vive solo	640	64
vive solo	360	36

Variable	Mean	Std Dev	Minimum	Maximum	N	Lower Quartile	Median	Upper Quartile
mes	20.9	12.1	4	72	1000	12	18	24
dm	3271.3	2822.8	250	18424	1000	1365	2319.5	3972.5







Esempio: ricodifica variabili

y	Frequency	Percent
buen	700	70
mal	300	30

cuenta	Frequency	Percent
bad running	332	33.2
good running	394	39.4
no	274	27.4

sexo	Frequency	Percent
hombre	598	59.8
mujer	402	40.2

uso	Frequency	Percent
privado	657	65.7
profesional	343	34.3

ppag	Frequency	Percent
pre buen pag	911	91.1
pre mal paga	89	8.9

estc	Frequency	Percent
no vive solo	640	64
vive solo	360	36

y2	1	mal
	0	buen

cuenta1	1	bad running
	0	otherwise

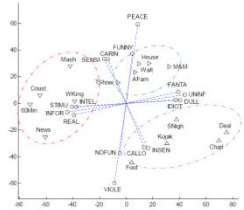
cuenta2	1	good running
	0	otherwise

estc1	1	no vive solo
	0	vive solo

ppag1	1	pre buen pagador
	0	pre mal pagador

sexo1	1	hombre
	0	mujer

uso1	1	privado
	0	profesional



Modello logit di regressione multipla

Per studiare la relazione che intercorre tra una variabile dicotomica e *k* variabili esplicative possiamo estendere il modello LOGIT visto precedentemente come

Equazione

$$\vartheta_i = \frac{\exp(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_K x_{iK})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_K x_{iK})}$$

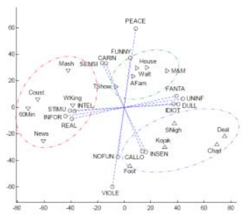
Assunzioni

$y_i|\boldsymbol{x}_i \sim \mathrm{Be}(\vartheta_i)$, indip.

o equivalentemente

$$\mathrm{logit}(\vartheta_i) = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_K x_{iK}$$

$y_i|\boldsymbol{x}_i \sim \mathrm{Be}(\vartheta_i)$, indip.



Modello lineare di utilità

Supponiamo sia vero il seguente modello lineare che lega la variabile risposta *U* (ad es. rischiosità di un individuo) ad alcune variabili esplicative ***x***

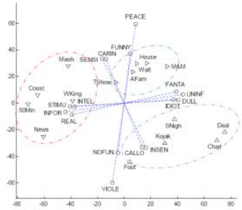
$$u_i = \beta_1 x_{i1} + \ldots + \beta_K x_{iK} + \varepsilon_i$$

Non riusciamo ad osservare completamente *U,* osserviamo solo se questa supera la soglia - β_0 (molto rischioso quindi default).

Per la probabilità di default risulta

$$\begin{aligned} 1 - \vartheta_i &= \Pr(u_i < -\beta_0) \\ &= \Pr(\beta_1 x_{i1} + \ldots + \beta_K x_{iK} + \varepsilon_i < -\beta_0) \\ &= \Pr(\varepsilon_i < -\beta_0 - \beta_1 x_{i1} + \ldots - \beta_K x_{iK}) \\ &= F_{cum}(-\beta_0 - \beta_1 x_{i1} + \ldots - \beta_K x_{iK}) \\ &= F_{cum}(-\mathbf{x}_i' \boldsymbol{\beta}) \end{aligned}$$

dove F_{cum} è la funzione di ripartizione della v.a. ε.



Modello logit

Una delle assunzioni maggiormente utilizzate per la distribuzione di ε è la **logistica**

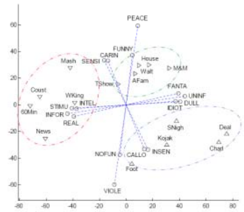
$$f(x) = \frac{\exp(x)}{[1+\exp(x)]^2}, F_{cum}(x) = \frac{\exp(x)}{1+\exp(x)}, E(X) = 0, V(X) = \frac{\pi^2}{3}$$

Sostituendo otteniamo

$$1 - \vartheta_i = F_{cum}(-\mathbf{x}_i'\boldsymbol{\beta}) = \frac{\exp(-\mathbf{x}_i'\boldsymbol{\beta})}{1+\exp(-\mathbf{x}_i'\boldsymbol{\beta})} \Rightarrow \vartheta_i = 1 - \frac{\exp(-\mathbf{x}_i'\boldsymbol{\beta})}{1+\exp(-\mathbf{x}_i'\boldsymbol{\beta})}$$

$$\Rightarrow \vartheta_i = \frac{1}{1+\exp(-\mathbf{x}_i'\boldsymbol{\beta})} \Rightarrow \vartheta_i = \frac{\exp(\mathbf{x}_i'\boldsymbol{\beta})}{1+\exp(\mathbf{x}_i'\boldsymbol{\beta})}$$

il modello **Logit**



Modello probit

l'altra assunzione molto utilizzata per la distribuzione di ε è la **normale standard**

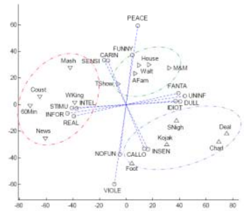
$$f(x)=\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}x^2\right\},F_{cum}(x)=\Phi(x),E(X)=0,V(X)=1$$

Sostituendo otteniamo

$$1-\vartheta_i=F_{cum}(-\mathbf{x}_i'\boldsymbol{\beta})=\Phi(-\mathbf{x}_i'\boldsymbol{\beta})$$
$$\Rightarrow\vartheta_i=1-\Phi(-\mathbf{x}_i'\boldsymbol{\beta})\Rightarrow\vartheta_i=\Phi(\mathbf{x}_i'\boldsymbol{\beta})$$

il modello **Probit**

I due modelli non sono molto diversi in quanto le due distribuzioni sono abbastanza somiglianti.

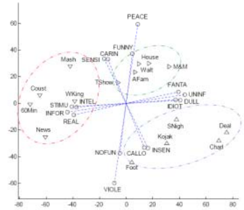


Esempio: stima

Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	204.382	8	<.0001
Score	190.987	8	<.0001
Wald	152.622	8	<.0001

Model Fit Statistics		
Criterion	**Intercept Only**	**Intercept and Covariates**
AIC	1223.729	1035.347
SC	1228.636	1079.517
-2 Log L	1221.729	1017.347

Analysis of Maximum Likelihood Estimates							
Parameter	**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**	**95% Confidence Limits**	
Intercept	1	0.67	0.3199	4.387	0.0362	0.043	1.297
cuenta1	1	-0.6346	0.1764	12.943	0.0003	-0.9804	-0.2889
cuenta2	1	-1.9517	0.2061	89.710	<.0001	-2.3555	-1.5478
dm	1	0.000032	0.000033	0.945	0.3309	-0.00003	0.000098
estc1	1	-0.3854	0.2194	3.087	0.0789	-0.8154	0.0445
mes	1	0.035	0.00785	19.914	<.0001	0.0196	0.0504
ppag1	1	-0.9884	0.253	15.268	<.0001	-1.4841	-0.4926
sexo1	1	-0.2235	0.2208	1.025	0.3115	-0.6563	0.2093
uso1	1	-0.4744	0.1605	8.740	0.0031	-0.7889	-0.1599



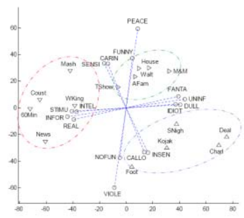
logit: test di Wald e LR

Test z (significatività delle singole variabili, può essere utilizzato per la backward selection)

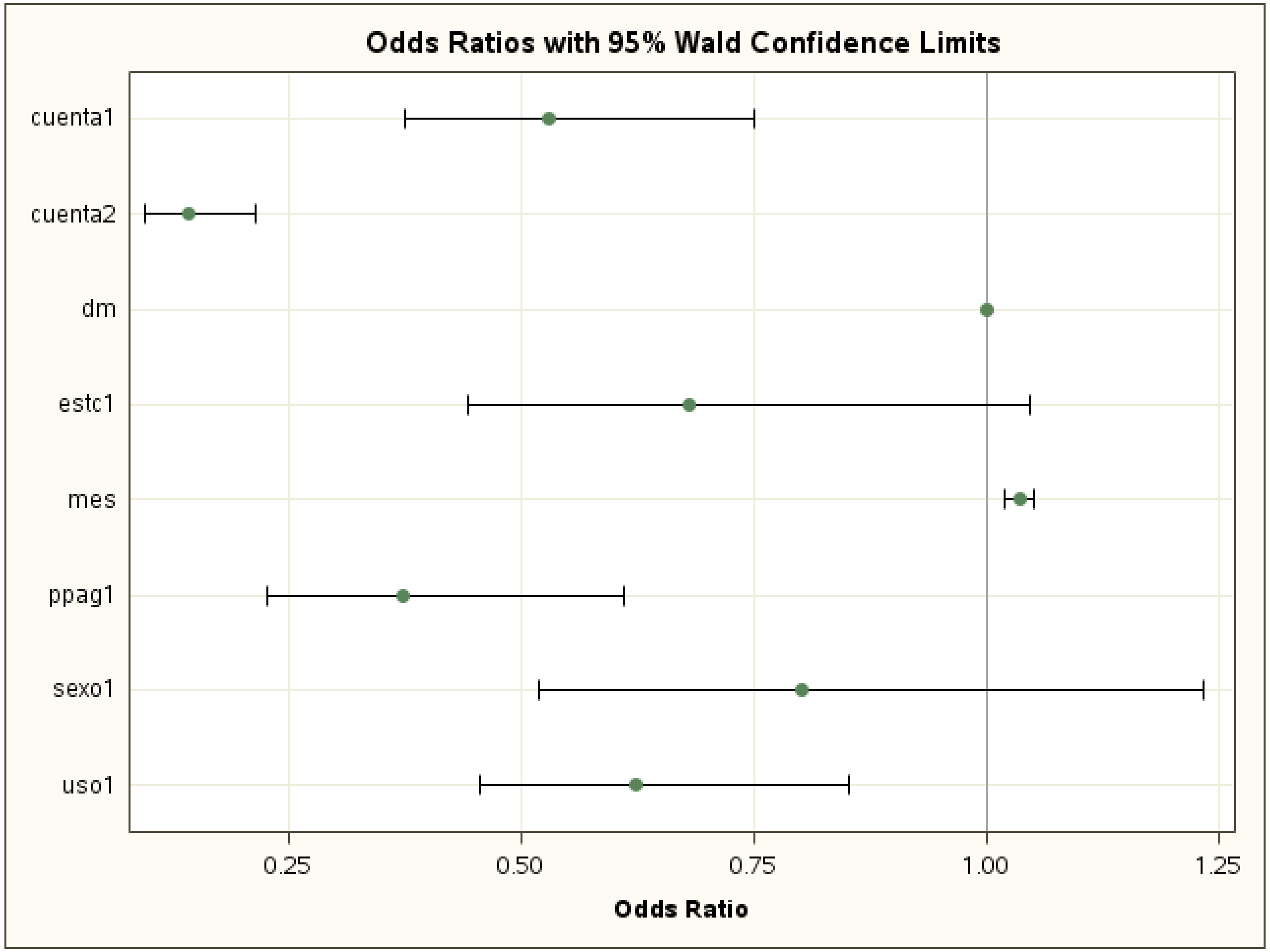
- Ipotesi nulla H_0: $b_j=0$
- Stat. test $w_{oss}= (\text{Coef.}/(\text{Std.Err.}))^2$
  $w_{oss} \sim \chi_1$ se H_0 vera e n suff. elevato
- regola rifiuto $w_{oss} > \chi_{1,\alpha}$
- p-value $\Pr\{W > w_{oss}\}$

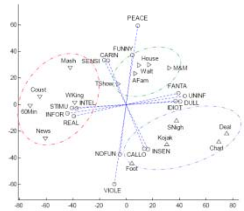
Test LR (significatività del modello nel complesso)

- Ipotesi nulla H_0: $b_1=b_2=\ldots=b_k=0$
- Stat. Test $LR_{oss}= 2Ll - 2Ll(H_0)$
  $LR_{oss} \sim \chi_k$ se H_0 vera e n suff. elevato
- regola rifiuto $LR_{oss} > \chi_{k,\alpha}$
- p-value $\Pr\{\chi_k > LR_{oss}\}$



logit: stime degli odds





logit: bontà della classificazione

Per accertare la bontà della classificazione, costruiamo una tabella di contingenza incrociando la *Y* osservata con quella teorica ($\hat{Y} = 1 \text{ se } \hat{\vartheta}_i > \text{cut off}$)

Cut-off = 0.5

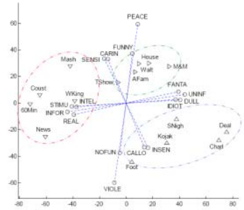
	$Y = d$	$Y = nd$	
$\hat{Y} = d$	**0.118**[1]	0.063	0.181
$\hat{Y} = nd$	0.182	**0.637**[1]	0.819
	0.3	0.7	1

Distribuzione congiunta $(\hat{Y}, Y)$

Una prima misura di bontà di classificazione è

1) ***Rate of correct classification* (*RCC*).** Numero casi correttamente classificati / *n*

Nell'esempio abbiamo RCC = 0.118 + 0.637 = 0.755



logit bc: sensitivity - specificity

Nel caso di eventi rari la RCC può assumere valori elevati anche per modelli poco abili nella classificazione. Se ad es. i default sono solo il 5% un modello banale che predice sempre il non default avrebbe RCC = 0.95. Per evitare questo problema si calcola (evento = default)

2) ***Sensitivity*** (abilità nel predire un *evento* correttamente) quota di *eventi* predetti come *eventi*.

3) ***Specificity*** (abilità nel predire un *non evento* correttamente) quota di *non eventi* predetti come *non eventi*.

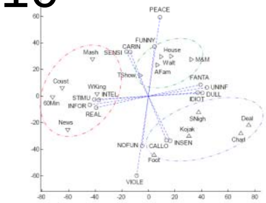
cut-off = 0.5

	$Y = d$	$Y = nd$
$\hat{Y} = d$	**0.393**[2]	0.090
$\hat{Y} = nd$	0.607	**0.910**[3]
	1	1

Tabella profili colonna

Distribuzioni di $\hat{Y} \mid Y$

Sensitivity = 0.118/0.300 = 0.393, Specificity = 0.637/0.700 = 0.910



Logit bc: false rates

In alternativa a Sensitivity e Specificity possiamo anche calcolare (evento = default)

4) ***False positive rate*** quota di casi predetti come *eventi* ma osservati come *non eventi*.

5) ***False negative rate*** quota di casi predetti come *non eventi* ma osservati come *eventi*.

cut-off = 0.5

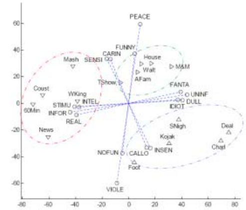
	$Y = d$	$Y = nd$	
$\hat{Y} = d$	0.652	**0.348**[4]	1
$\hat{Y} = nd$	**0.222**[5]	0.778	1

Tabella profili riga
Distribuzioni di $Y \mid \hat{Y}$

False positive rate = 0.063/0.181 = 0.348

False negative rate = 0.182/0.819 = 0.222

Non sono misure di bontà. La classificazione è ottima se sono 0.

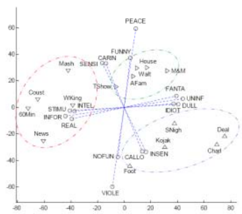


Esempio: classificazione

Classification Table									
Prob Level	Correct		Incorrect		Percentages				
	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.5	113	632	68	187	74.5	37.7	90.3	37.6	22.8

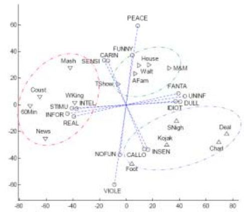
	Y=1	Y=0	
Ystim=1	113	68	181
Ystim=0	187	632	819
	300	700	1000

Tasso corretta classificazione =(113+632)/100=74,5
Sensitivity = (113/300)*100=37,7%
Specificity = (632/700)*100=90,3%
Falso positivo = (68/181)*100=37,6%
Falso negativo = (187/819)*100=22,8%

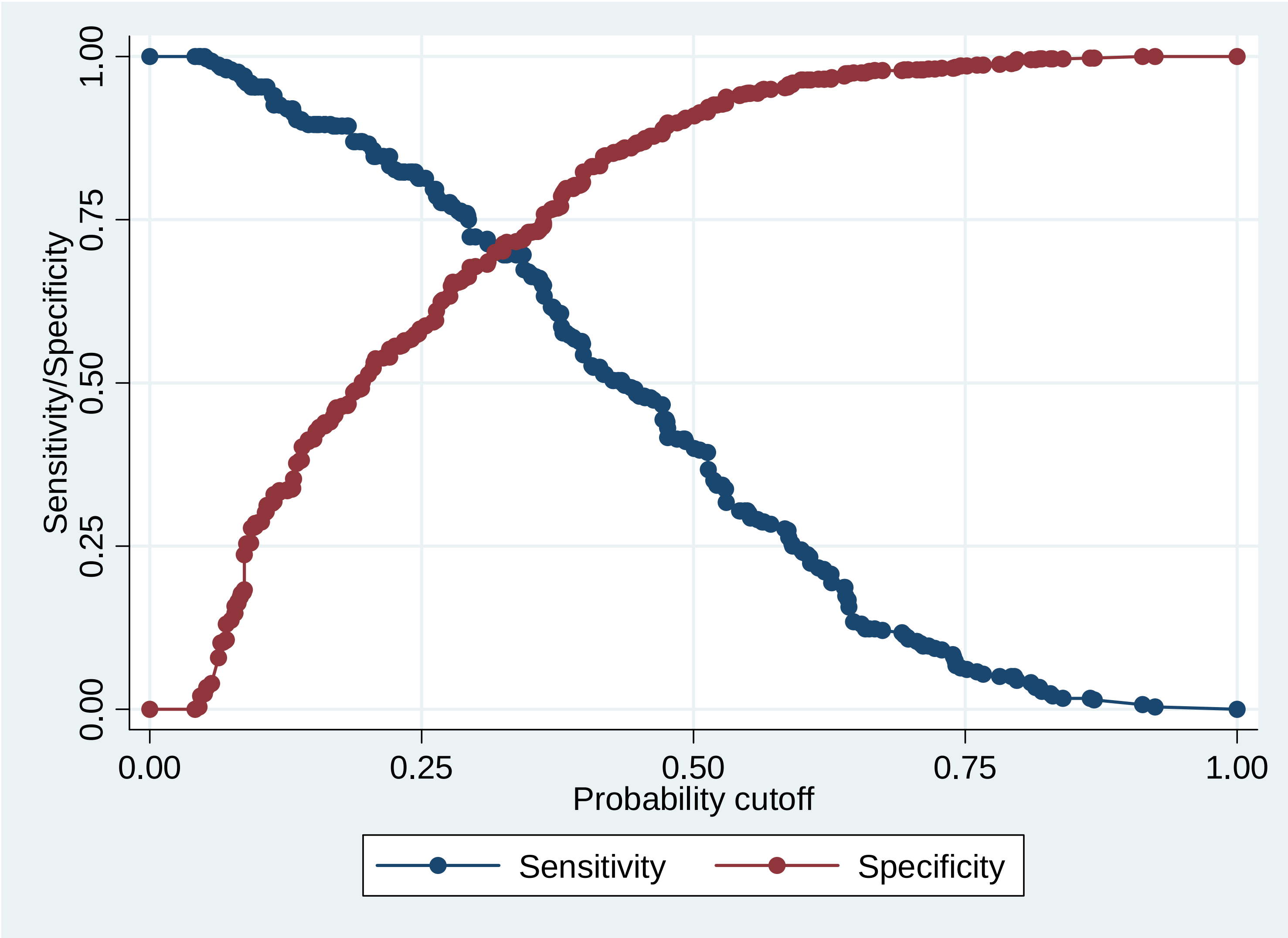


logit bc: scelta cut-off

- Un cut-off sulla scala delle probabilità può essere tradotto nella scala logit. Ad esempio, un livello di probabilità uguale a 0.5 implica la soglia log(0.5/(1-0.5)) = 0. La regola di classificazione diventa: *i* è predetta come evento se $\beta_1 x_{i1} + \ldots + \beta_K x_{iK} > -\beta_0$.
- Le misure precedentemente introdotte dipendono dal cut-off scelto. Per questo è importante studiare come cambiano al variare della soglia scelta. Allo scopo si rappresentano graficamente le due curve **SS**. Una formata dai punti (sensitivity,cut-off) e l’altra dai punti (specificity,cut-off).
- Il cut-off ottimale viene deciso sulla base dei diversi costi derivanti da una errata predizione di un default e/o non default.

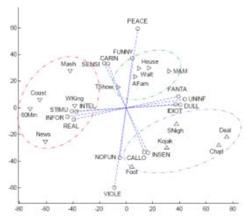


curve ss


Sensitivity/Specificity
0.00
0.25
0.50
0.75
1.00
0.00
0.25
0.50
0.75
1.00
Probability cutoff
Sensitivity
Specificity

logit bc: curva ROC

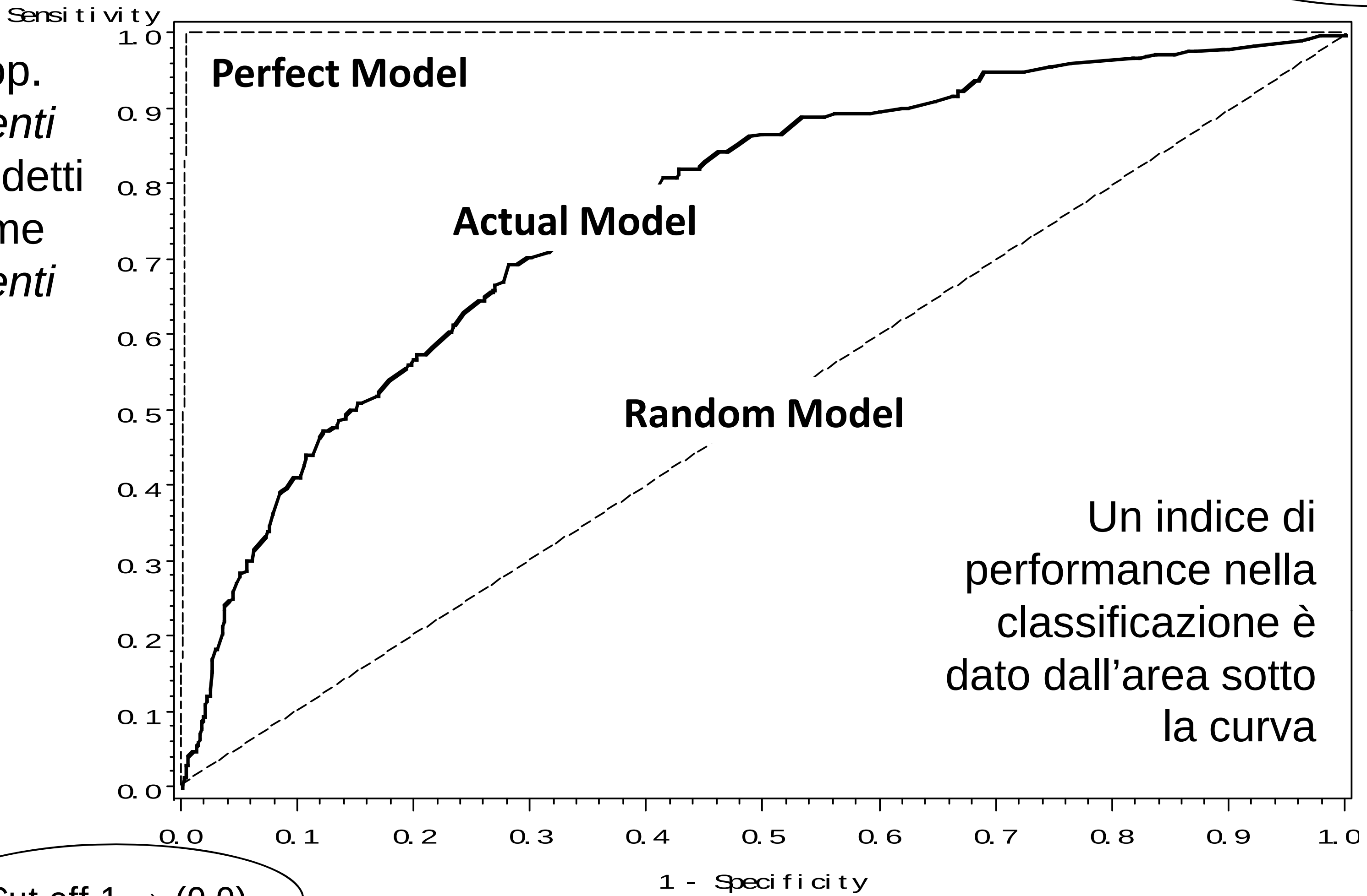
- Le variazioni in termini di sensitivity e specificity corrispondenti a diversi livelli di cut-off possono anche essere rappresentate mediante la curva **ROC**.
- La curva ROC è costituita da tutti i punti di coordinate (sensitivity,1-specificity) che si creano considerando tutti i possibili valori di cut-off tra 0 e 1.
- L'area sottesa dalla curva è un indice di bontà della classificazione che varia tra 0 e 1.
- La curva ROC è' molto utile per confrontare tra loro diversi modelli di classificazione.



Curva ROC

Cut off 0 ⇒ (1,1)

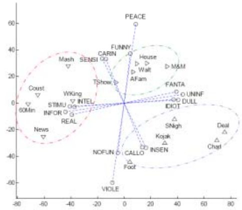
Prop. *eventi* predetti come *eventi*



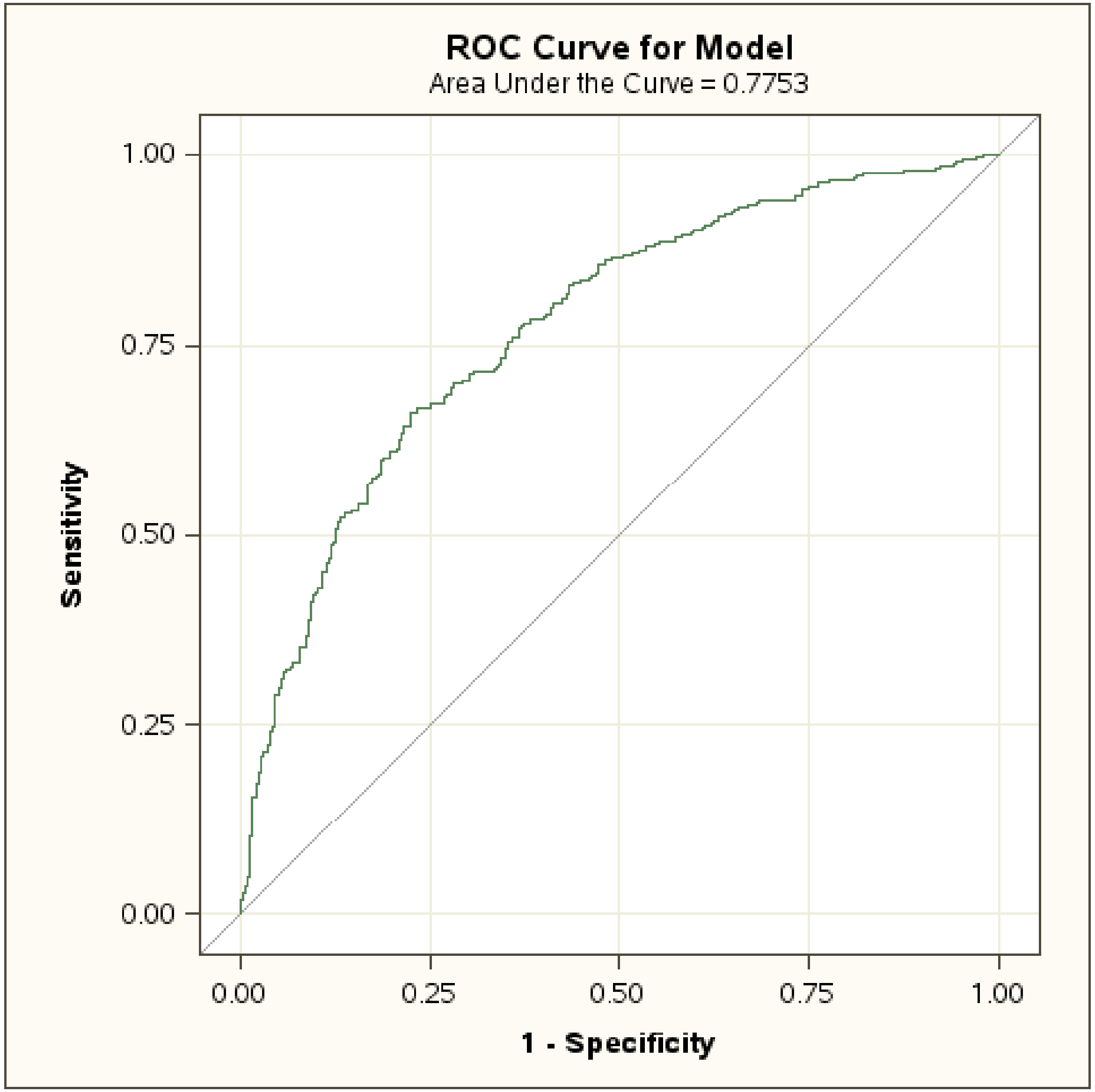
Un indice di performance nella classificazione è dato dall'area sotto la curva

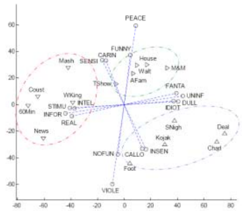
Cut off 1 ⇒ (0,0)

Prop. *non eventi* predetti come *eventi*



Esempio: curva ROC

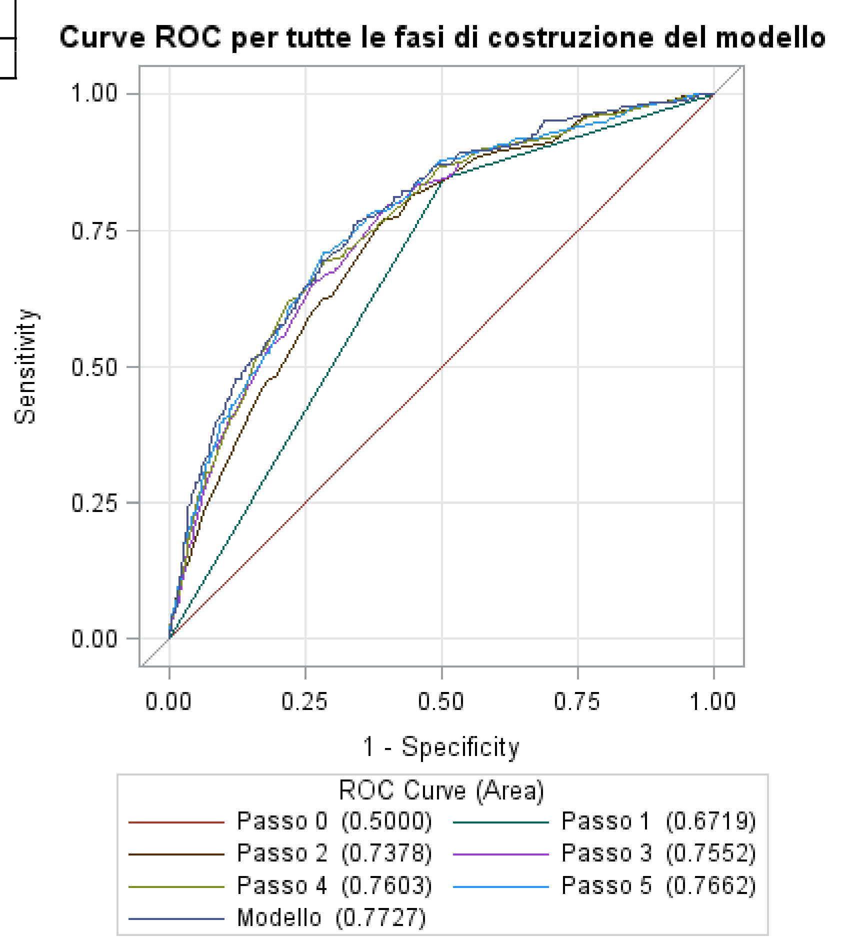




Esempio: Forward selection

Summary of Forward Selection						
Step	Effect Entered	DF	Number In	Score Chi-Square	Pr > ChiQuadr	Variable Label
1	cuenta2	1	1	103.9648	<.0001	cuenta==good running
2	mes	1	2	40.952	<.0001	mes
3	ppag1	1	3	18.8406	<.0001	ppag==pre buen pagador
4	cuenta1	1	4	11.7336	0.0006	cuenta==bad running
5	estc1	1	5	10.1391	0.0015	estc==no vive solo
6	uso1	1	6	8.706	0.0032	uso==privado

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiQuadr
Intercept	1	0.6437	0.3179	4.0987	0.0429
cuenta1	1	-0.6174	0.1757	12.3417	0.0004
cuenta2	1	-1.9377	0.2055	88.9342	<.0001
estc1	1	-0.5327	0.1591	11.2081	0.0008
mes	1	0.0389	0.00629	38.1125	<.0001
ppag1	1	-0.9877	0.2527	15.2807	<.0001
uso1	1	-0.4694	0.1597	8.6408	0.0033



logit bc: out of sample

- Nella trattazione precedente tutte le 5 misure sono state calcolate in sample, i.e. sui dati utilizzati per stimare il modello. Questo comporta una sovrastima dell'abilità del modello nella classificazione.
- Si può ovviare al problema spezzando in due parti il campione: training e test. Sulla parte training si stima il modello, le stime così ottenute vengono utilizzate per classificare le osservazioni della parte test del campione e calcolare le 5 misure.