COURS « TRAITEMENTS STATISTIQUES DANS L’UNIVERS BANCAIRE »

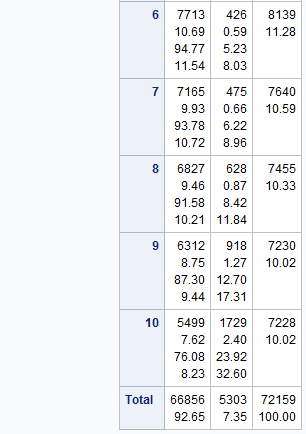
SECTION 8 – EVALUATION DU MODELE

Cette section décrit les techniques d’évaluation d’un modèle lors de sa construction

1. **EVALUER LA PERFORMANCE**

Le découpage de l’échantillon de construction en quantiles (ici en déciles) permet de s’assurer de la performance du score dans le classement des défauts :





Exemple de programmation pour obtenir ce tableau :

**proc** **logistic** descending

data = lpdm.score\_29\_logist

outest = scolog1 ;

model defaut = score\_var\_D: score\_var\_E: score\_var\_F: score\_var\_G:

score\_var\_H: score\_var\_I:

/ selection = none ctable link = logit ;

output out = scolog2 prob = proba xbeta = estim ;

**run** ;

**proc** **score** data = lpdm.score\_29\_logist

score = scolog1

out = ficsco2

type = PARMS ;

var score\_var\_D: score\_var\_E: score\_var\_F: score\_var\_G: score\_var\_H:

score\_var\_I: ;

**run** ;

**data** ficsco2 ;

set ficsco2 ;

rename defaut2 = note\_new ;

bidon = **1** ;

**run** ;

**proc** **univariate** data = ficsco2 ;

var note\_new ;

output out = ficsco3

pctlpts = **0** **10** **20** **30** **40** **50** **60** **70** **80** **90** **100**

pctlpre = quant ;

**run** ;

**data** ficsco3 ;

set ficsco3 ;

bidon = **1** ;

**run** ;

**data** ficsco2 ; merge ficsco2 ficsco3 ;

by bidon ;

trnote = (**1** \* ( note\_new lt quant10 ) )

+ (**2** \* ( quant10 le note\_new lt quant20 ) )

+ (**3** \* ( quant20 le note\_new lt quant30 ) )

+ (**4** \* ( quant30 le note\_new lt quant40 ) )

+ (**5** \* ( quant40 le note\_new lt quant50 ) )

+ (**6** \* ( quant50 le note\_new lt quant60 ) )

+ (**7** \* ( quant60 le note\_new lt quant70 ) )

+ (**8** \* ( quant70 le note\_new lt quant80 ) )

+ (**9** \* ( quant80 le note\_new lt quant90 ) )

+ (**10** \* ( quant90 le note\_new ) ) ;

**run** ;

**proc** **freq** data = ficsco2 ;

tables trnote \* defaut ;

title 'Performance du score par déciles' ;

**run** ;

On peut également tracer la courbe de performance du score : proportion du nombre de défauts en fonction de la proportion du nombre de bons clients, après classement par note de score décroissante.

Courbe de performance d'un score :

On classe les dossiers par note de score décroissante (ou croissante si le score modélise la probabilité de non défaut). On représente en abscisse la proportion de bons et en ordonnée la proportion de mauvais dossiers correspondante.



La courbe se lit ainsi :

En isolant 20 % des bons dossiers (les 20 % ayant les meilleures (resp. moins bonnes) notes de score), on parvient à isoler 82 % des dossiers mauvais.

Un bon score de risque présente ce type de pouvoir discriminant (10/60 et 20/80).

1. **EVALUER LA PERFORMANCE ET LA STABILITE SUR L’ECHANTILLON DE VALIDATION**

Le même travail doit être effectué sur l’échantillon de validation (et le cas échéant sur l’échantillon hors temps) :

* On réapplique le modèle (la grille obtenue) aux données de cet échantillon
* On réapplique le même découpage en quantiles (attention : il ne s’agit pas de définir de nouveaux quantiles sur l’échantillon de validation, mais bien de reproduire le découpage en quantiles opéré sur l’écha²ntillon de construction)

On peut alors comparer :

* Les taux de défaut observés par quantile d’origine entre les deux échantillons
* La stabilité du découpage par note de score (en mesurant la stabilité des quantiles au travers d’un indice de stabilité).