

# ETUDE DE CAS EN ASSURANCE

## ANALYSE DISCRIMINANTE

### SCORE - SEGMENTATION



Pages 1 et 2: Présentation des données

Pages 3 à 5: Caractérisation d'une variable qualitative

Pages 6 à 8: Analyse des correspondances multiples

Pages 9 à 17: Analyse discriminante sur les facteurs de l'ACM

Pages 18 à 21: Création et étude d'une fonction de score

Pages 22 à 27 : Procédure LOGISTIC SAS

Pages 28 à 29 : Comparaison de scores

# Étude de cas ‘Assurance’ (d’après une étude de Philippe Périé, CISIA)

## *I.1. Les données*

L’échantillon est constitué de 1106 assurés Belges observés en 1992 et répartis en 2 groupes.

- **les assurés qui n’ont eu aucun accident dans l’année**, qui sont au nombre de 556.
- **les assurés ayant eu au moins un sinistre dans l’année**; ils sont au nombre de 550.

On entend par assuré une personne physique ou morale.

Les variables de départ, concernent le souscripteur et le véhicule :

- Le sexe (3 modalités)
- Le code usage (2 modalités)
- L’année de construction (2 modalités)
- L’année de naissance (9 modalités)
- Le degré de bonus-malus de l’année et celui de l’année précédente (12 modalités)
- La date d’effet de la police (4 modalités)
- Le code postal recodé par arrondissement et regroupé par proximité et caractère urbain, suburbain ou rural (12 modalités)
- La puissance (14 modalités)
- La marque (17 modalités)
- Le code fractionnement des quittances (4 modalités)
- La valeur catalogue (7 modalités)
- La présence ou non de garantie DM, vol ou autre (2 modalités pour chaque variable)
- Le nombre d’accidents dans l’année transformé en variable à 2 modalités (pas d’accidents et au moins 1 accident dans l’année écoulée)

## ***1.2 Étude préalable, sélection des variables***

Les variables originelles sont au nombre de 19 : la SINISTRALITE et 18 variables dont 16 nominales et 2 continues. Une première étape nous a permis de créer les variables 20 à 30 (caractérisations automatiques, analyses factorielles, tests de regroupements, recodages de continues en qualitatives, essais de création de variables d'interaction ...).

Dictionnaire originel des données :

---

1 . Sinistralité RC - SNB11 2-3	( 2 MODALITES )
2 . Code usage - CUSA 5-6	( 2 MODALITES )
3 . Age de l'assuré - DNAI 8-9	( 9 MODALITES )
4 . Sexe - SEXE 11-12	( 3 MODALITES )
5 . Code Langue - CLAN 14-15	( 2 MODALITES )
6 . Code postal souscripteur - POSS2 17-18	( 12 MODALITES )
7 . Bonus-malus RC(année en cours) - GBMC 20-21	( 11 MODALITES )
8 . Bonus-Malus Année -1 - GBM1	( 11 MODALITES )
9 . Date effet police - DPEP 26-27	( 5 MODALITES )
10 . Marque/type du véhicule - MARQ2 29-30	( 19 MODALITES )
11 . Puissance du véhicule - PUIS 32-33	( 14 MODALITES )
12 . Valeur catalogue - VAL3 35-36	( 8 MODALITES )
13 . Année de construction du véhicule - DCOS 38-39	( 6 MODALITES )
14 . Présence garantie 2 - GAR2 41-42	( 2 MODALITES )
15 . Presence garantie 3 - GAR3 44-45	( 2 MODALITES )
16 . Primes Acquises RC 1991 en francs belges - TPA11	( CONTINUE )
17 . Charge Sinistre RC 1991 en francs belges - CHA11	( CONTINUE )
18 . Primes acquises RC 1991 en francs belges - TP11	( 8 MODALITES )
19 . Charge Sinistre RC 1991 en francs belges - CHA11	( 8 MODALITES )

Variables ajoutées :

20 . Age de l'assuré (8 mod) - DNAI 8-9	( 8 MODALITES )
21 . Bonus-malus RC(année en cours) - GBMC 20-21	( 9 MODALITES )
22 . Bonus-Malus Année -1 (9 mod) - GBM1	( 9 MODALITES )
23 . Puissance du véhicule (12 mod) - PUIS 32-33	( 12 MODALITES )
24 . Age de l'assuré (3 mod) - DNAI 8-9	( 3 MODALITES )
25 . Code postal souscripteur (2 mod) - POSS2 17-18	( 2 MODALITES )
26 . Bonus-malus Année -1 (2 mod) - GBM1	( 2 MODALITES )
27 . Date effet Police (2 mod) - DPEP 26-27	( 2 MODALITES )
28 . Puissance du véhicule (2 mod) - PUIS 32-33	( 2 MODALITES )
29 . Année de construction du véhicule - DCOS 38-39	( 2 MODALITES )
30 . Primes acquises RC 1991 en francs belges - TP11	( 3 MODALITES )

---

Choix des variables discriminantes : Caractérisation des 2 groupes '0 sinistre' et '1 sinistre ou plus' par les variables nominales (proc DEMOD de SPAD)

**Classe: 0 sin**

Libellés des variables	Modalités caractéristiques	% de la modalité dans l'échantillon	% de la classe dans la modalité	Valeur-Test	Poids
Bonus-malus Année -1 (2 mod) - GBM1	B-M 1 (-1)	49.64	83.97	23.19	549
Date effet Police (2 mod) - DPEP 26-27	<86 Police	56.87	68.20	13.85	629
Age de l'assuré (3 mod) - DNAI 8-9	Naiss ???	44.85	68.95	11.26	496
Année de construction du véhicule (2 mod) - DCOS 38-39	33-89 DCOS	74.41	59.05	10.06	823
Code postal souscripteur (2 mod) - POSS2 17-18	Autres codes	66.82	60.35	9.55	739
Code usage - CUSA 5-6	Privé	83.27	55.48	7.80	921
Code Langue - CLAN 14-15	Lang néerland.	25.50	63.48	5.09	282
Puissance du véhicule (2 mod) - PUIS 32-33	10-39 Puis	19.62	64.06	4.48	217
Sexe - SEXE 11-12	masculin	71.16	54.38	4.24	787
Puissance du véhicule (2 mod) - PUIS 32-33	40-349 Puis	80.38	46.91	-4.48	889
Code Langue - CLAN 14-15	Lang franç.	74.50	45.75	-5.09	824
Sexe - SEXE 11-12	sexe autre	6.33	20.00	-5.26	70
Code usage - CUSA 5-6	Profess.	16.73	24.32	-7.80	185
Code postal souscripteur (2 mod) - POSS2 17-18	Bruxelles	33.18	29.97	-9.55	367
Année de construction du véhicule (2 mod) - DCOS 38-39	90-91 DCOS	25.59	24.74	-10.06	283
Date effet Police (2 mod) - DPEP 26-27	autres polices	43.13	26.62	-13.85	477
Age de l'assuré (3 mod) - DNAI 8-9	1950-73	27.94	17.48	-13.94	309
Bonus-malus Année -1 (2 mod) - GBM1	Autres B-M (-1)	50.36	17.06	-23.19	557

**Classe: > 1 sin**

<b>Libellés des variables</b>	<b>Modalités caractéristiques</b>	<b>% de la modalité dans l'échantillon</b>	<b>% de la classe dans la modalité</b>	<b>Valeur-Test</b>	<b>Poids</b>
Bonus-malus Année -1 (2 mod) - GBM1	Autres B-M (-1)	50.36	82.94	23.19	557
Age de l'assuré (3 mod) - DNAI 8-9	1950-73	27.94	82.52	13.94	309
Date effet Police (2 mod) - DPEP 26-27	autres polices	43.13	73.38	13.85	477
Année de construction du véhicule (2 mod) - DCOS 38-39	90-91 DCOS	25.59	75.27	10.06	283
Code postal souscripteur (2 mod) - POSS2 17-18	Bruxelles	33.18	70.03	9.55	367
Code usage - CUSA 5-6	Profess.	16.73	75.68	7.80	185
Sexe - SEXE 11-12	sexe autre	6.33	80.00	5.26	70
Code Langue - CLAN 14-15	Lang franç.	74.50	54.25	5.09	824
Puissance du véhicule (2 mod) - PUIS 32-33	40-349 Puis	80.38	53.09	4.48	889
Sexe - SEXE 11-12	masculin	71.16	45.62	-4.24	787
Puissance du véhicule (2 mod) - PUIS 32-33	10-39 Puis	19.62	35.94	-4.48	217
Code Langue - CLAN 14-15	Lang néerland.	25.50	36.52	-5.09	282
Code usage - CUSA 5-6	Privé	83.27	44.52	-7.80	921
Code postal souscripteur (2 mod) - POSS2 17-18	Autres codes	66.82	39.65	-9.55	739
Année de construction du véhicule (2 mod) - DCOS 38-39	33-89 DCOS	74.41	40.95	-10.06	823
Age de l'assuré (3 mod) - DNAI 8-9	Naiss ???	44.85	31.05	-11.26	496
Date effet Police (2 mod) - DPEP 26-27	<86 Police	56.87	31.80	-13.85	629
Bonus-malus Année -1 (2 mod) - GBM1	B-M 1 (-1)	49.64	16.03	-23.19	549

Croisements de 'sinistralité' avec les autres variables (proc DEMOD de SPAD)

<b>Libellé de la variable</b>	<b>Khi-2</b>	<b>Nb. de degrés de liberté</b>	<b>Effectifs théoriques inférieur à 5</b>	<b>Valeur-Test</b>	<b>Probabilité</b>
Age de l'assuré (3 mod) - DNAI 8-9	203.18	2	0	99.99	0.000
Bonus-malus Année -1 (2 mod) - GBM1	492.54	1	0	99.99	0.000
Date effet Police (2 mod) - DPEP 26-27	185.94	1	0	13.59	0.000
Année de construction du véhicule (2 mod) - DCOS 38-39	97.84	1	0	9.82	0.000
Code postal souscripteur (2 mod) - POSS2 17-18	89.32	1	0	9.38	0.000
Code usage - CUSA 5-6	58.59	1	0	7.57	0.000
Sexe - SEXE 11-12	32.99	2	0	5.27	0.000
Code Langue - CLAN 14-15	25.69	1	0	4.94	0.000
Puissance du véhicule (2 mod) - PUIS 32-33	19.84	1	0	4.30	0.000

### ***I.3 Analyse des correspondances multiples (proc CORMU de SPAD)***

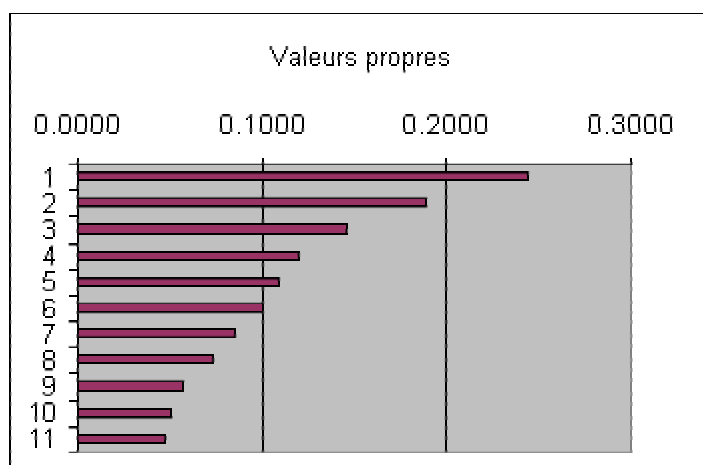
Nous réalisons une analyse des correspondances multiples sur les variables explicatives en mettant en illustrative la variable 'SINIST'. De cette façon, nous pourrons visualiser les modalités et les conjonctions de modalités qui déterminent l'appartenance des individus à tel ou tel groupe.

#### Valeurs propres

**Tableau des valeurs propres**

**Trace de la matrice: 1.22222**

Numéro	Valeur propre	Pourcentage	Pourcentage cumulé
1	0.2438	19.95	19.95
2	0.1893	15.49	35.44
3	0.1457	11.92	47.36
4	0.1201	9.82	57.18
5	0.1091	8.92	66.11
6	0.0999	8.17	74.28
7	0.0855	7.00	81.28
8	0.0732	5.99	87.26
9	0.0573	4.68	91.95
10	0.0511	4.18	96.13
11	0.0473	3.87	100.00



## Coordonnées et valeurs test des modalités actives axes 1 à 5

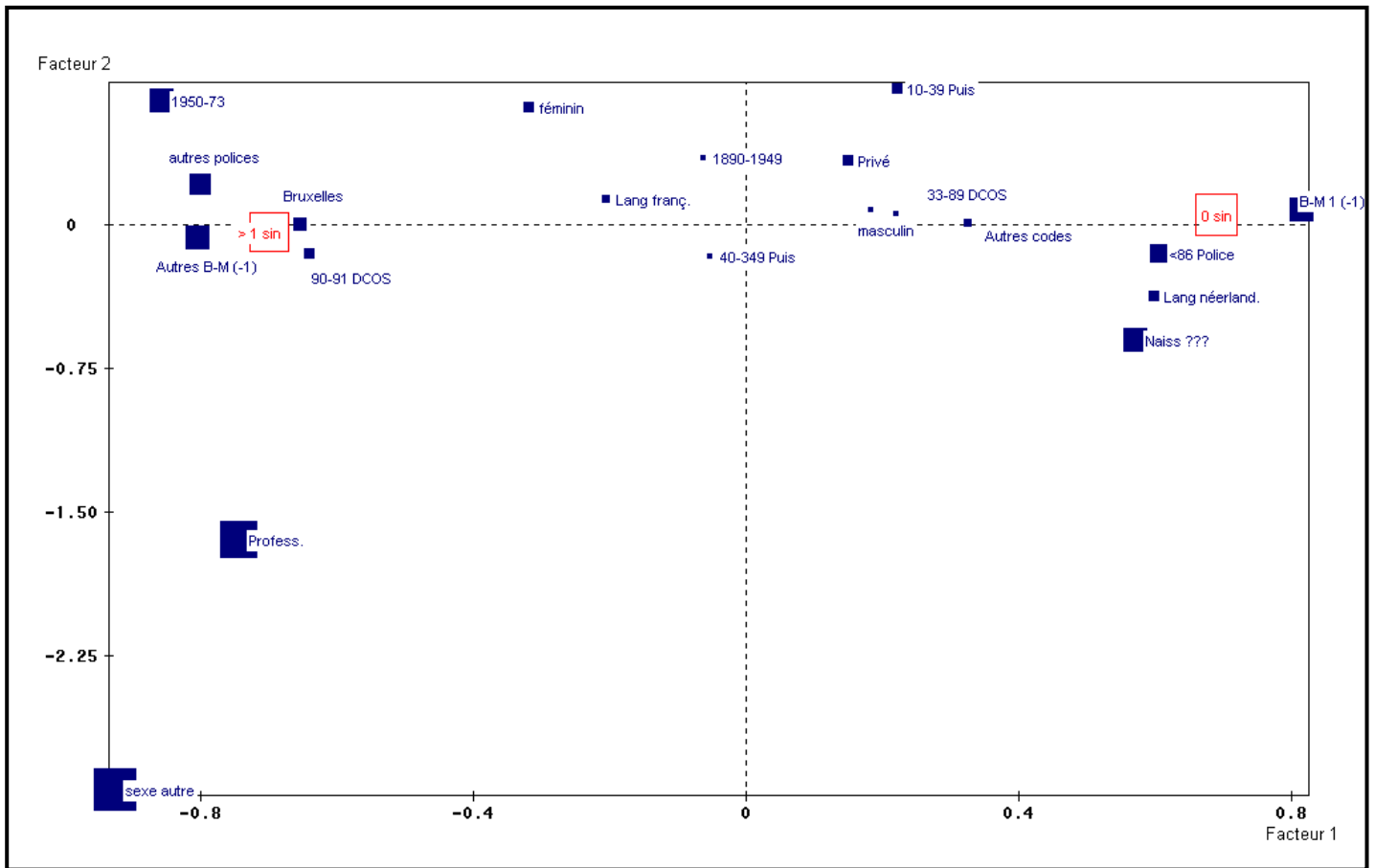
MODALITES			VALEURS-TEST					COORDONNEES					DISTO.
IDEN - LIBELLE	EFF.	P.ABS	1	2	3	4	5	1	2	3	4	5	
2 . Code usage - CUSA 5-6													
CUS1 - Profess.	185	185.00	-11.1	-24.5	-1.5	2.1	3.7	-0.74	-1.64	-0.10	0.14	0.25	4.98
CUS2 - Privé	921	921.00	11.1	24.5	1.5	-2.1	-3.7	0.15	0.33	0.02	-0.03	-0.05	0.20
4 . Sexe - SEXE 11-12													
MASC - masculin	787	787.00	9.5	3.7	17.4	-9.6	-16.6	0.18	0.07	0.33	-0.18	-0.32	0.41
FEMI - féminin	249	249.00	-5.7	10.8	-15.6	8.5	16.2	-0.32	0.61	-0.87	0.47	0.90	3.44
SOCI - sexe autre	70	70.00	-8.0	-25.5	-5.7	3.4	3.1	-0.93	-2.95	-0.66	0.39	0.36	14.80
5 . Code Langue - CLAN 14-15													
FRAN - Lang franç.	824	824.00	-11.6	7.4	-14.5	-17.9	1.1	-0.20	0.13	-0.26	-0.31	0.02	0.34
NEER - Lang néerland.	282	282.00	11.6	-7.4	14.5	17.9	-1.1	0.60	-0.38	0.75	0.92	-0.06	2.92
24 . Age de l'assuré (3 mod) - DNAI 8-9													
AGE1 - 1890-1949	301	301.00	-1.3	6.9	13.9	-16.7	15.1	-0.06	0.34	0.69	-0.82	0.74	2.67
AGE2 - 1950-73	309	309.00	-17.7	13.2	-0.2	13.8	-13.2	-0.86	0.64	-0.01	0.67	-0.64	2.58
AGE? - Naiss ???	496	496.00	17.1	-18.2	-12.3	2.5	-1.6	0.57	-0.61	-0.41	0.08	-0.05	1.23
25 . Code postal souscripteur (2 mod) - POSS2 17-18													
COD1 - Bruxelles	367	367.00	-15.3	-0.1	-15.8	-14.7	-7.1	-0.65	0.00	-0.67	-0.63	-0.30	2.01
COD2 - Autres codes	739	739.00	15.3	0.1	15.8	14.7	7.1	0.32	0.00	0.33	0.31	0.15	0.50
26 . Bonus-malus Année -1 (2 mod) - GBM1													
BM01 - B-M 1 (-1)	549	549.00	26.9	2.3	-1.1	-3.4	5.9	0.82	0.07	-0.03	-0.10	0.18	1.01
BM02 - Autres B-M (-1)	557	557.00	-26.9	-2.3	1.1	3.4	-5.9	-0.80	-0.07	0.03	0.10	-0.18	0.99
27 . Date effet Police (2 mod) - DPEP 26-27													
P86 - <86 Police	629	629.00	23.2	-5.9	-10.8	-7.0	-0.9	0.61	-0.15	-0.28	-0.18	-0.02	0.76
P87 - autres polices	477	477.00	-23.2	5.9	10.8	7.0	0.9	-0.80	0.20	0.37	0.24	0.03	1.32
28 . Puissance du véhicule (2 mod) - PUIS 32-33													
PU01 - 10-39 Puis	217	217.00	3.6	11.5	-15.5	12.3	9.1	0.22	0.70	-0.94	0.75	0.55	4.10
PU04 - 40-349 Puis	889	889.00	-3.6	-11.5	15.5	-12.3	-9.1	-0.05	-0.17	0.23	-0.18	-0.13	0.24
29 . Année de construction du véhicule (2 mod) - DCOS 38-39													
DC01 - 33-89 DCOS	823	823.00	12.5	3.0	-9.9	2.0	-17.9	0.22	0.05	-0.17	0.04	-0.32	0.34
DC02 - 90-91 DCOS	283	283.00	-12.5	-3.0	9.9	-2.0	17.9	-0.64	-0.15	0.51	-0.10	0.92	2.91

## Coordonnées et valeurs test des modalités illustratives axes 1 à 5

1 . Sinistralité RC - SNB11 2-3													
SRC0 - 0 sin	556	556.00	23.1	1.5	-2.3	-1.5	2.8	0.69	0.05	-0.07	-0.04	0.08	0.99
SRC1 - > 1 sin	550	550.00	-23.1	-1.5	2.3	1.5	-2.8	-0.70	-0.05	0.07	0.04	-0.08	1.01



# Plan Factoriel 1, 2



## ***I.4 DISQUAL : Analyse discriminante sur les facteurs issus de l'ACM (proc dis2gfm de SPAD)***

A/ Modèle 'Complet'.

<b>Libellé de la variable</b>	<b>Khi-2</b>	<b>Nb. de degrés de liberté</b>	<b>Effectifs théoriques inférieur à 5</b>	<b>Valeur-Test</b>
Age de l'assuré (3 mod) - DNAI 8-9	203.18	2	0	99.99
Bonus-malus Année -1 (2 mod) - GBM1	492.54	1	0	99.99
Date effet Police (2 mod) - DPEP 26-27	185.94	1	0	13.59
Année de construction du véhicule (2 mod) - DCOS 38-39	97.84	1	0	9.82
Code postal souscripteur (2 mod) - POSS2 17-18	89.32	1	0	9.38
Code usage - CUSA 5-6	58.59	1	0	7.57
Sexe - SEXE 11-12	32.99	2	0	5.27
Code Langue - CLAN 14-15	25.69	1	0	4.94
Puissance du véhicule (2 mod) - PUIS 32-33	19.84	1	0	4.30

Échantillon de Base : 742 individus (67%)

Échantillon test : 364 individus

## Fonction linéaire discriminante

FACTEURS ..... NUM IDEN	CORRELATIONS FACTEURS AVEC F.L.D. (SEUIL= 0.07)	COEFFICIENTS FONCTION DISC.	REGRESSION	ECARTS TYPES (RES. TYPE REG.)	T STUDENT	PROBA
1 F 1	0.719	6.9064	1.4569	0.0485	30.06	0.000
2 F 2	0.055	0.7149	0.1508	0.0554	2.72	0.007
3 F 3	-0.078	-0.8211	-0.1732	0.0630	2.75	0.006
4 F 4	-0.030	-0.4615	-0.0973	0.0685	1.42	0.155
5 F 5	0.083	1.2581	0.2654	0.0729	3.64	0.000
6 F 6	0.064	1.0274	0.2167	0.0748	2.90	0.004
7 F 7	-0.001	0.2169	0.0458	0.0828	0.55	0.581
8 F 8	0.090	1.3133	0.2770	0.0888	3.12	0.002
9 F 9	-0.074	-1.1383	-0.2401	0.1028	2.34	0.020
10 F 10	-0.150	-3.3193	-0.7002	0.1079	6.49	0.000
11 F 11	-0.056	-1.4830	-0.3128	0.1107	2.83	0.005
CONSTANTE		0.093575	0.016617	0.0240	0.6910	0.4898
.....						
R2 =	0.57923	F =	91.35686	PROBA =	0.000	
D2 =	5.49176	T2 =	1018.69159	PROBA =	0.000	
.....						

## Fonction linéaire de Fisher reconstituée à partir des variables d'origine (Fonction de score 'brute')

FONCTION LINEAIRE DE FISHER RECONSTITUEE A PARTIR DES VARIABLES D'ORIGINE						
VARIABLES ..... NUM IDEN LIBELLE	COEFFICIENTS FONCTION DISC.	REGRESSION	ECARTS TYPES (RES. TYPE REG.)	T STUDENT	PROBA	
2 CUS1 Profess.	-5.5612	-1.1731	0.5818	2.02	0.044	
2 CUS2 Privé	1.1171	0.2356	0.1169	2.02	0.044	
4 MASC masculin	0.6335	0.1336	0.1453	0.92	0.358	
4 FEMI féminin	1.8990	0.4006	0.4193	0.96	0.340	
4 SOCI sexe autre	-13.8768	-2.9273	1.0011	2.92	0.004	
5 FRAN Lang franç.	-0.5336	-0.1126	0.1344	0.84	0.403	
5 NEER Lang néerland.	1.5591	0.3289	0.3927	0.84	0.403	
24 AGE1 1890-1949	1.0782	0.2274	0.3642	0.62	0.533	
24 AGE2 1950-73	-13.1694	-2.7781	0.4124	6.74	0.000	
24 AGE? Naiss ???	7.5500	1.5927	0.2968	5.37	0.000	
25 COD1 Bruxelles	-6.7782	-1.4298	0.3360	4.26	0.000	
25 COD2 Autres codes	3.3661	0.7101	0.1669	4.26	0.000	
26 BM01 B-M 1 (-1)	21.0068	4.4314	0.2812	15.76	0.000	
26 BM02 Autres B-M (-1)	-20.7051	-4.3677	0.2771	15.76	0.000	
27 P86 <86 Police	2.2157	0.4674	0.2327	2.01	0.045	
27 P87 autres polices	-2.9218	-0.6164	0.3068	2.01	0.045	
28 PU01 10-39 Puis	5.4794	1.1559	0.4537	2.55	0.011	
28 PU04 40-349 Puis	-1.3375	-0.2821	0.1107	2.55	0.011	
29 DC01 33-89 DCOS	3.9748	0.8385	0.1311	6.39	0.000	
29 DC02 90-91 DCOS	-11.5591	-2.4384	0.3813	6.39	0.000	
CONSTANTE	0.093575	0.016617				
.....						

*Tableau de classement de l'échantillon d'apprentissage*

RESULTATS DE LA DISCRIMINATION LINEAIRE DE FISHER  
TABLEAU DES EFFECTIFS DES GROUPES

GROUPES D'ORIGINE	GROUPES D'AFFECTATION	
	SRC0	SRC1
SRC0	314	59
SRC1	34	335

TABLEAU DE CLASSEMENT

GROUPES D'ORIGINE	POURCENTAGES DES CLASSEMENTS		
	BIEN CLASSES	MAL CLASSES	TOTAL
SRC0	314.00 ( 84.18)	59.00 ( 15.82)	373.00 (100.00)
SRC1	335.00 ( 90.79)	34.00 ( 9.21)	369.00 (100.00)
TOTAL	649.00 ( 87.47)	93.00 ( 12.53)	742.00 (100.00)

*ADL sur l'échantillon test : tableaux de classement*

ANALYSE DISCRIMINANTE LINEAIRE SUR L'ECHANTILLON : TEST  
ENTRE LES 2 GROUPES : 0 sin ET > 1 sin  
VARIABLE DE GROUPE NUMERO 1 : Sinistralité RC - SNB11 2-3  
RESULTATS DE LA DISCRIMINATION LINEAIRE DE FISHER

TABLEAU DES EFFECTIFS DES GROUPES

GROUPES D'ORIGINE	GROUPES D'AFFECTATION	
	SRC0	SRC1
SRC0	152	31
SRC1	32	149

TABLEAU DE CLASSEMENT

GROUPES D'ORIGINE	POURCENTAGES DES CLASSEMENTS		
	BIEN CLASSES	MAL CLASSES	TOTAL
SRC0	152.00 ( 83.06)	31.00 ( 16.94)	183.00 (100.00)
SRC1	149.00 ( 82.32)	32.00 ( 17.68)	181.00 (100.00)
TOTAL	301.00 ( 82.69)	63.00 ( 17.31)	364.00 (100.00)

## B/ Modèle réduit 1 : sélection des axes factoriels

Pour la sélection, nous ne gardons que les T de Student élevés sur les axes du modèle complet. Cette sélection peut se faire directement puisque les axes factoriels sont des variables orthogonales.

FONCTION LINEAIRE DISCRIMINANTE						
FACTEURS	CORRELATIONS	COEFFICIENTS		ECARTS	T	PROBA
.....	FACTEURS	FONCTION	REGRESSION	TYPES	STUDENT	
NUM IDEN	AVEC F.L.D.	DISC.		(RES. TYPE REG.)		
	(SEUIL= 0.07)					
1 F 1	0.719	6.9064	1.4569	0.0485	30.06	0.000
2 F 2	0.055	0.7149	0.1508	0.0554	2.72	0.007
3 F 3	-0.078	-0.8211	-0.1732	0.0630	2.75	0.006
4 F 4	-0.030	-0.4615	-0.0973	0.0685	1.42	0.155
5 F 5	0.083	1.2581	0.2654	0.0729	3.64	0.000
6 F 6	0.064	1.0274	0.2167	0.0748	2.90	0.004
7 F 7	-0.001	0.2169	0.0458	0.0828	0.55	0.581
8 F 8	0.090	1.3133	0.2770	0.0888	3.12	0.002
9 F 9	-0.074	-1.1383	-0.2401	0.1028	2.34	0.020
10 F 10	-0.150	-3.3193	-0.7002	0.1079	6.49	0.000
11 F 11	-0.056	-1.4830	-0.3128	0.1107	2.83	0.005
CONSTANTE		0.093575	0.016617	0.0240	0.6910	0.4898
.....						
R2 =	0.57923	F =	91.35686	PROBA =	0.000	
D2 =	5.49176	T2 =	1018.69159	PROBA =	0.000	
.....						

## Fonction linéaire de Fisher reconstituée à partir des variables d'origine

FONCTION LINEAIRE DE FISHER RECONSTITUEE A PARTIR DES VARIABLES D'ORIGINE						
VARIABLES	COEFFICIENTS		ECARTS	T	PROBA	
.....	FONCTION	REGRESSION	TYPES	STUDENT		
NUM IDEN LIBELLE	DISC.		(RES. TYPE REG.)			
2 CUS1 Profess.	-5.2399	-1.1089	0.5775	1.92	0.055	
2 CUS2 Privé	1.0525	0.2227	0.1160	1.92	0.055	
4 MASC masculin	0.6239	0.1320	0.1094	1.21	0.228	
4 FEMI féminin	1.7831	0.3774	0.3006	1.26	0.210	
4 SOCI sexe autre	-13.3575	-2.8269	0.9981	2.83	0.005	
5 FRAN Lang franç.	-0.8021	-0.1698	0.1056	1.61	0.108	
5 NEER Lang néerland.	2.3438	0.4960	0.3086	1.61	0.108	
24 AGE1 1890-1949	-0.1780	-0.0377	0.3180	0.12	0.906	
24 AGE2 1950-73	-12.0564	-2.5515	0.3896	6.55	0.000	
24 AGE? Naiss ???	7.6190	1.6124	0.2954	5.46	0.000	
25 COD1 Bruxelles	-7.7786	-1.6462	0.3037	5.42	0.000	
25 COD2 Autres codes	3.8630	0.8175	0.1508	5.42	0.000	
26 BM01 B-M 1 (-1)	20.8088	4.4038	0.2804	15.70	0.000	
26 BM02 Autres B-M (-1)	-20.5100	-4.3406	0.2764	15.70	0.000	
27 P86 <86 Police	1.9845	0.4200	0.2298	1.83	0.068	
27 P87 autres polices	-2.6169	-0.5538	0.3031	1.83	0.068	
28 PU01 10-39 Puis	7.1745	1.5184	0.3309	4.59	0.000	
28 PU04 40-349 Puis	-1.7513	-0.3706	0.0808	4.59	0.000	
29 DC01 33-89 DCOS	3.8924	0.8238	0.1224	6.73	0.000	
29 DC02 90-91 DCOS	-11.3196	-2.3956	0.3560	6.73	0.000	
CONSTANTE	0.097630	0.017546				
.....						

### *ADL sur l'échantillon de base : tableaux de classement*

ANALYSE DISCRIMINANTE LINEAIRE SUR L'ECHANTILLON : DE BASE  
ENTRE LES 2 GROUPEs : 0 sin ET > 1 sin  
VARIABLE DE GROUPE NUMERO 1 : Sinistralité RC - SNB11 2-3  
RESULTATS DE LA DISCRIMINATION LINEAIRE DE FISHER  
TABLEAU DES EFFECTIFS DES GROUPEs

GROUPEs D'ORIGINE	GROUPEs D'AFFECTION	
	SRC0	SRC1
SRC0	313	60
SRC1	33	336

#### TABLEAU DE CLASSEMENT

GROUPEs D'ORIGINE	POURCENTAGES DES CLASSEMENTS		TOTAL
	BIEN CLASSEs	MAL CLASSEs	
SRC0	313.00 ( 83.91)	60.00 ( 16.09)	373.00 (100.00)
SRC1	336.00 ( 91.06)	33.00 ( 8.94)	369.00 (100.00)
TOTAL	649.00 ( 87.47)	93.00 ( 12.53)	742.00 (100.00)

### *ADL sur l'échantillon test : tableaux de classement*

ANALYSE DISCRIMINANTE LINEAIRE SUR L'ECHANTILLON : TEST  
ENTRE LES 2 GROUPEs : 0 sin ET > 1 sin  
VARIABLE DE GROUPE NUMERO 1 : Sinistralité RC - SNB11 2-3  
RESULTATS DE LA DISCRIMINATION LINEAIRE DE FISHER  
TABLEAU DES EFFECTIFS DES GROUPEs

GROUPEs D'ORIGINE	GROUPEs D'AFFECTION	
	SRC0	SRC1
SRC0	152	31
SRC1	32	149

#### TABLEAU DE CLASSEMENT

GROUPEs D'ORIGINE	POURCENTAGES DES CLASSEMENTS		TOTAL
	BIEN CLASSEs	MAL CLASSEs	
SRC0	152.00 ( 83.06)	31.00 ( 16.94)	183.00 (100.00)
SRC1	149.00 ( 82.32)	32.00 ( 17.68)	181.00 (100.00)
TOTAL	301.00 ( 82.69)	63.00 ( 17.31)	364.00 (100.00)

C/ Modèle réduit 2 : Analyse sur les variables les plus fortement liées à la sinistralité et sélection sur les axes factoriels issus de l'ACM

Libellé de la variable	Khi-2	Nb. de degrés de liberté	Effectifs théoriques inférieur à 5	Valeur-Test
Age de l'assuré (3 mod) - DNAI 8-9	203.18	2	0	99.99
Bonus-malus Année -1 (2 mod) - GBM1	492.54	1	0	99.99
Date effet Police (2 mod) - DPEP 26-27	185.94	1	0	13.59
Année de construction du véhicule (2 mod) - DCOS 38-39	97.84	1	0	9.82
Code postal souscripteur (2 mod) - POSS2 17-18	89.32	1	0	9.38
Code usage - CUSA 5-6	58.59	1	0	7.57
Sexe - SEXE 11-12	32.99	2	0	5.27
Code Langue - CLAN 14-15	25.69	1	0	4.94
Puissance du véhicule (2 mod) - PUIS 32-33	19.84	1	0	4.30

**FONCTION LINEAIRE DISCRIMINANTE SUR LES AXES RETENUS**

FACTEURS ..... NUM IDEN	CORRELATIONS FACTEURS AVEC F.L.D. (SEUIL= 0.07)	COEFFICIENTS FONCTION DISC.	REGRESSION	ECARTS TYPES (RES. TYPE REG.)	T STUDENT	PROBA
1 F 1	0.714	5.6085	1.2124	0.0410	29.59	0.000
2 F 2	0.143	1.5432	0.3336	0.0554	6.02	0.000
3 F 5	0.067	0.8129	0.1757	0.0642	2.74	0.006
4 F 6	-0.159	-2.6238	-0.5672	0.0860	6.60	0.000
5 F 7	0.075	1.5185	0.3283	0.0902	3.64	0.000
CONSTANTE		0.100417	0.018641	0.0242	0.7698	0.4417
.....						
R2 =	0.56881	F =	194.18361	PROBA =	0.000	
D2 =	5.26266	T2 =	976.19495	PROBA =	0.000	
.....						

**FONCTION LINEAIRE DE FISHER RECONSTITUEE A PARTIR DES VARIABLES D'ORIGINE**

VARIABLES		COEFFICIENTS		ECARTS	T	PROBA
NUM	IDEN LIBELLE	FONCTION DISC.	REGRESSION	TYPES (RES. TYPE REG.)	STUDENT	
2	CUS1 Profess.	-6.6745	-1.4429	0.3396	4.25	0.000
2	CUS2 Privé	1.3407	0.2898	0.0682	4.25	0.000
24	AGE1 1890-1949	-0.3353	-0.0725	0.1326	0.55	0.585
24	AGE2 1950-73	-6.8141	-1.4730	0.2248	6.55	0.000
24	AGE? Naiss ???	4.4486	0.9617	0.1904	5.05	0.000
25	COD1 Bruxelles	-4.9880	-1.0783	0.1428	7.55	0.000
25	COD2 Autres codes	2.4771	0.5355	0.0709	7.55	0.000
26	BM01 B-M 1 (-1)	13.9444	3.0144	0.1843	16.35	0.000
26	BM02 Autres B-M (-1)	-13.7441	-2.9711	0.1817	16.35	0.000
27	P86 <86 Police	1.5569	0.3366	0.1527	2.20	0.028
27	P87 autres polices	-2.0530	-0.4438	0.2014	2.20	0.028
29	DC01 33-89 DCOS	2.8785	0.6223	0.0687	9.06	0.000
29	DC02 90-91 DCOS	-8.3711	-1.8096	0.1997	9.06	0.000
	CONSTANTE	0.100417	0.018641			

*ADL sur l'échantillon de base : tableaux de classement*

TABLEAU DES EFFECTIFS DES GROUPES

GROUPES D'ORIGINE	GROUPES D'AFFECTATION	
	SRC0	SRC1
SRC0	308	65
SRC1	40	329

TABLEAU DE CLASSEMENT

GROUPES D'ORIGINE	POURCENTAGES DES CLASSEMENTS		
	BIEN CLASSES	MAL CLASSES	TOTAL
SRC0	308.00 ( 82.57)	65.00 ( 17.43)	373.00 (100.00)
SRC1	329.00 ( 89.16)	40.00 ( 10.84)	369.00 (100.00)
TOTAL	637.00 ( 85.85)	105.00 ( 14.15)	742.00 (100.00)



## ANALYSE DISCRIMINANTE LINEAIRE SUR L'ECHANTILLON : TEST

### TABLEAU DES EFFECTIFS DES GROUPES

GROUPES D'ORIGINE	GROUPES D'AFFECTATION	
	SRC0	SRC1
SRC0	151	32
SRC1	36	145

### TABLEAU DE CLASSEMENT

GROUPES D'ORIGINE	POURCENTAGES DES CLASSEMENTS		
	BIEN CLASSES	MAL CLASSES	TOTAL
SRC0	151.00 ( 82.51)	32.00 ( 17.49)	183.00 (100.00)
SRC1	145.00 ( 80.11)	36.00 ( 19.89)	181.00 (100.00)
TOTAL	296.00 ( 81.32)	68.00 ( 18.68)	364.00 (100.00)

## ESTIMATIONS BOOTSTRAP DES EFFECTIFS ET POURCENTAGES: 1000 TIRAGES ALEATOIRES

### ENTRE ETOILES : ECART-TYPE DES POURCENTAGES

GROUPES	RAPPEL DU CALCUL DE BASE		BOOTSTRAP		
	EFFECTIFS ET (POURCENTAGES)		EFFECTIFS ET (POURCENTAGES)		TOTAL
	BIEN CLASSES	MAL CLASSES	BIEN CLASSES	MAL CLASSES	
SRC0	308.00 ( 82.57)	65.00 ( 17.43)	307.38 ( 82.41) * 1.95*	65.62 ( 17.59) * 1.95*	373.00 (100.00)
SRC1	329.00 ( 89.16)	40.00 ( 10.84)	328.15 ( 88.93) * 1.46*	40.85 ( 11.07) * 1.46*	369.00 (100.00)
TOTAL	637.00 ( 85.85)	105.00 ( 14.15)	635.53 ( 85.65) * 1.23*	106.47 ( 14.35) * 1.23*	742.00 (100.00)

## ESTIMATIONS BOOTSTRAP POUR LA FONCTION DISCRIMINANTE

FACTEURS		CORRELATIONS		COEFFICIENTS		
NUMERO	NOM	FACTEURS		FONCTION		
		F.L.D.		DISCRIMINANTE		
		MOYENNE	ECART-TYPE	MOYENNE	ECART-TYPE	MOY/E-T
1	F 1	0.715	0.021	5.708117	0.549869	10.3809
2	F 2	0.142	0.034	1.558156	0.294125	5.2976
3	F 5	0.066	0.037	0.813168	0.352014	2.3100
4	F 6	-0.159	0.037	-2.671309	0.564356	4.7334
5	F 7	0.075	0.035	1.561516	0.561031	2.7833
	CONSTANTE			0.104669	0.086503	1.2100

**ESTIMATIONS BOOTSTRAP POUR LA FONCTION LINEAIRE DE FISHER RECONSTITUEE A PARTIR DES VARIABLES D'ORIGINE**

VARIABLES		COEFFICIENTS		
.....		FONCTION		
NUM.	IDEN LIBELLE	DISCRIMINANTE		
		MOYENNE	ECART-TYPE	MOY/E-T
.....				
2	CUS1 Profess.	-6.736358	1.889277	3.5656
2	CUS2 Privé	1.353117	0.379496	3.5656
24	AGE1 1890-1949	-0.368415	0.684785	0.5380
24	AGE2 1950-73	-6.907277	1.385330	4.9860
24	AGE? Naiss ???	4.526706	1.097951	4.1229
25	COD1 Bruxelles	-5.035182	0.881501	5.7121
25	COD2 Autres codes	2.500558	0.437769	5.7121
26	BM01 B-M 1 (-1)	14.214258	1.769688	8.0321
26	BM02 Autres B-M (-1)	-14.010093	1.744270	8.0321
27	P86 <86 Police	1.603848	0.837824	1.9143
27	P87 autres polices	-2.114918	1.104803	1.9143
29	DC01 33-89 DCOS	2.913123	0.437435	6.6596
29	DC02 90-91 DCOS	-8.471754	1.272117	6.6596
	CONSTANTE	0.104669	0.086503	1.2100
.....				

## ***I.5 Création et étude d'une fonction de score (proc SCORE de SPAD)***

### Le principe

La procédure SCORE est exécutée après une analyse discriminante concernant deux groupes (proc DIS2G). Elle calcule une *fonction de score* qui est une modification de la fonction discriminante, destinée à faciliter son interprétation et son utilisation. On introduit en particulier des "zones" de décision (rouge, verte, centrale) à partir d'une *tolérance d'erreur de classement*, et on édite divers résultats permettant d'apprécier les performances de la discrimination. Dans la suite, les deux groupes à discriminer sont appelés:

- groupe des scores "*forts*" ou groupe "1"
- groupe des scores "*faibles*" ou groupe "2".

C'est l'utilisateur qui définit en entrée la correspondance entre les modalités 1 et 2 de la variable de groupe, et les attributs "*fort*" et "*faible*" des scores.

La fonction discriminante fournie par la procédure DIS2G doit avoir été calculée préalablement sur des variables nominales, c'est à dire en fait sur les axes factoriels résultant d'une analyse des correspondances multiples (analyse discriminante "qualitative"). Le passage des résultats entre DIS2G et SCORE est réalisé au moyen d'un fichier contenant les coefficients (fichier de type NCOEF).

L'utilisateur a cependant la possibilité d'introduire lui-même en paramètre l'ensemble des coefficients d'une fonction linéaire discriminante calculée par ailleurs (par exemple, des estimations "bootstrap").

La procédure SCORE transforme les coefficients en utilisant les deux règles suivantes:

- Coefficient minimum de chaque variable: pour chaque variable nominale, le plus petit coefficient est mis à la valeur 0. Ainsi le score minimum possible d'un individu est 0. Il est obtenu pour un individu qui, *pour chaque variable*, posséderait la modalité affectée du coefficient transformé 0.

- Maximum possible de la fonction score: la valeur du *score maximum* possible est un choix de l'utilisateur (par exemple 1000). Ce maximum correspond à la somme des coefficients transformés les plus grands dans chaque variable.

Le score attribué à un individu s'obtient en additionnant les coefficients transformés associés aux modalités de l'individu. La fonction de score transformée classe les individus de la même façon que la fonction discriminante initiale.

D'autre part la procédure SCORE peut être utilisée pour tester le comportement d'un jeu de coefficients transformés, modifiés manuellement par l'utilisateur. En général il s'agira des coefficients transformés fournis par une première exécution de la procédure, et "arrondis" à des valeurs proches pour en faciliter l'utilisation.

L'utilisateur peut définir un taux appelé "*tolérance d'erreur de classement*" noté TEC. Ce taux permet de calculer des zones sur l'échelle de la fonction de score:

- une **zone "verte"**, qui correspond à la région des scores élevés, où l'on s'attend à trouver la plupart des individus du groupe 1. Dans cette zone, un mal classé est un individu du groupe 2 affecté d'un score "trop" élevé. La frontière est calculée pour contenir un taux de mal classés ne dépassant pas TEC.

- une **zone "rouge"** du côté des scores faibles, contenant essentiellement des individus du groupe 2, donc bien classés, et un pourcentage ne dépassant pas TEC d'individus du groupe 1, donc mal classés.

- une **zone intermédiaire**, entre les frontières des zones rouge et verte, où l'affectation à un groupe est laissée indécise. Cette zone d'indécision diminue quand l'utilisateur augmente la tolérance de mauvais classement TEC.

**Attribution des scores faibles ou forts :**

ETUDE DE LA FONCTION SCORE

POUR LA VARIABLE DE GROUPE : Sinistralité RC - SNB11 2-3

	EFFECTIF
MODALITE 1 : " GROUPE DES SCORES FORTS " = 0 sin	371.
MODALITE 2 : " GROUPE DES SCORES FAIBLES " = > 1 sin	370.
ENSEMBLE	741.

**REMARQUE :**

*Dans tout ce qui suit, un nombre de point élevé indique une sinistralité plus faible. Le groupe des scores forts est celui des personnes n'ayant eu aucun sinistre.*

*La procédure 'SCORE' de SPAD permet à l'utilisateur de choisir la modalité qu'il va définir comme celle des scores forts, contrairement à la procédure 'DIS2G' qui réalise l'analyse discriminante.*

**Coefficients des fonctions discriminantes et score**

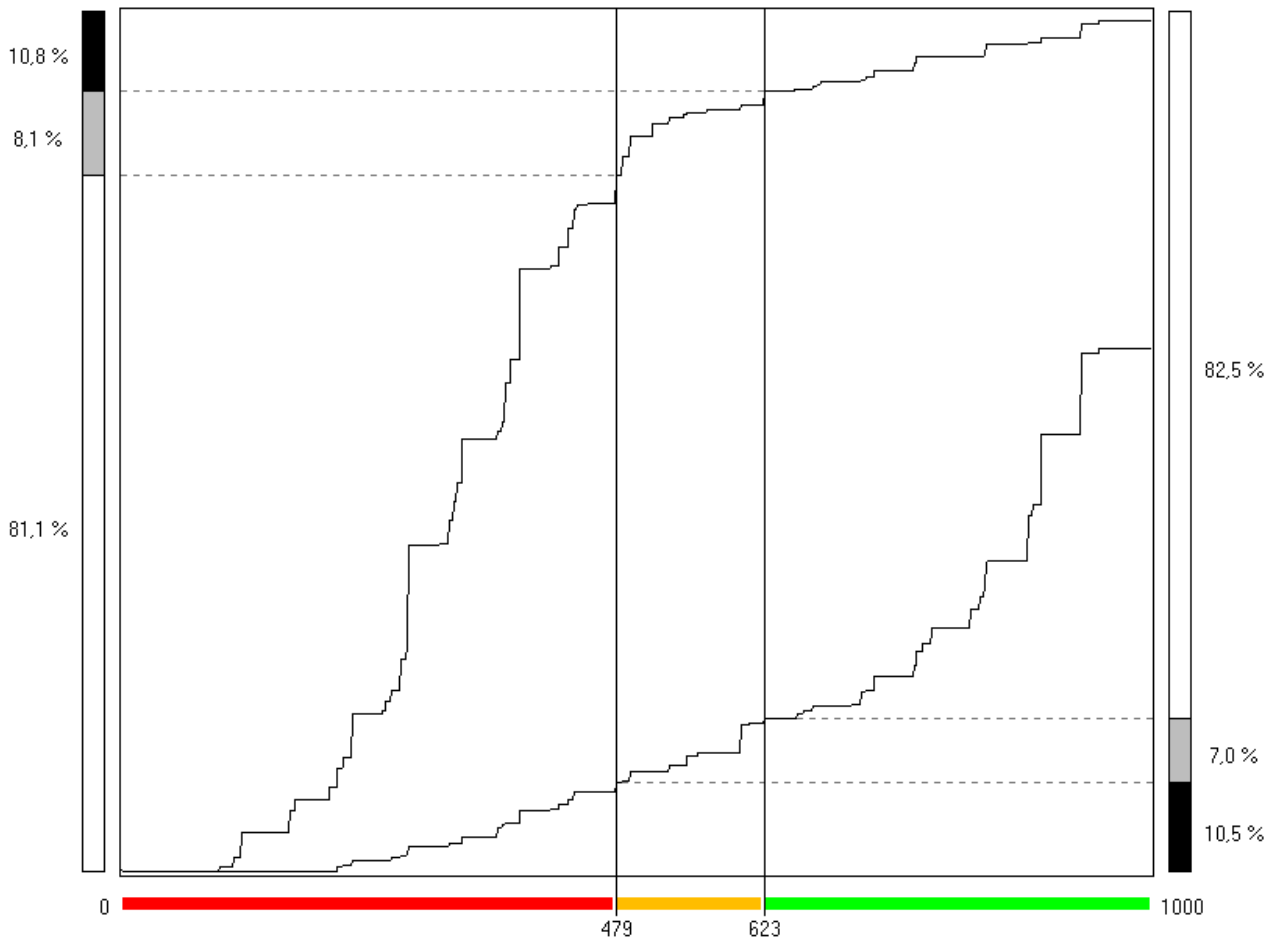
IDEN	LIBELLES	COEFFICIENTS FONCTION DISCRIMINANTE	COEFFICIENTS TRANSFORMES (SCORE)
2	. Code usage - CUSA 5-6		
prof	- Profess.	-6.675	0.00
priv	- Privé	1.341	115.68
24	. Age de l'assuré (3 mod) - DNAI 8-9		
AGE1	- 1890-1949	-0.335	93.50
AGE2	- 1950-73	-6.814	0.00
AGE?	- Naiss ???	4.449	162.54
25	. Code postal souscripteur (2 mod) - POSS2 17-18		
COD1	- Bruxelles	-4.988	0.00
COD2	- Autres codes	2.477	107.74
26	. Bonus-malus Année -1 (2 mod) - GBM1		
BM01	- B-M 1 (-1)	13.944	399.60
BM02	- Autres B-M (-1)	-13.744	0.00
27	. Date effet Police (2 mod) - DPEP 26-27		
P86	- <86 Police	1.557	52.10
P87	- autres polices	-2.053	0.00
29	. Année de construction du véhicule (2 mod) - DCOS 38-39		
DC01	- 33-89 DCOS	2.879	162.35
DC02	- 90-91 DCOS	-8.371	0.00

## Coefficients réordonnées de la fonction de score

IDEN	LIBELLES	COEFFICIENTS DU SCORE	HISTOGRAMMES DES POINTS DE SCORE
+-----+			
26	. Bonus-malus Année -1 (2 mod) - GBM1		
BM01	- B-M 1 (-1)	399.60	*****
BM02	- Autres B-M (-1)	0.00	*
+-----+			
24	. Age de l'assuré (3 mod) - DNAI 8-9		
AGE?	- Naiss ???	162.54	*****
AGE1	- 1890-1949	93.50	*****
AGE2	- 1950-73	0.00	*
+-----+			
29	. Année de construction du véhicule (2 mod) - DCOS 38-39		
DC01	- 33-89 DCOS	162.35	*****
DC02	- 90-91 DCOS	0.00	*
+-----+			
2	. Code usage - CUSA 5-6		
priv	- Privé	115.68	*****
prof	- Profess.	0.00	*
+-----+			
25	. Code postal souscripteur (2 mod) - POSS2 17-18		
COD2	- Autres codes	107.74	*****
COD1	- Bruxelles	0.00	*
+-----+			
27	. Date effet Police (2 mod) - DPEP 26-27		
P86	- <86 Police	52.10	***
P87	- autres polices	0.00	*
+-----+			

***Distribution de l'échantillon en fonction du score (sortie haute résolution)***

Limites pour un taux d'erreur de classement fixé à 10% : 479 et 623.



## I.6 Régression logistique (proc Logistic SAS v8)

### Le modèle, sélection par STEPWISE, représentation de la courbe ROC

La proc LOGISTIC de SAS version 8 permet de traiter les variables nominales telles quelles sans avoir à créer des colonnes d'indicatrices pour chacune des modalités.

Les options sont classiques :

- `descending` indique à SAS que la catégorie à modéliser est celle qui a la plus forte valeur (SINIST = 2 correspond à la catégorie des gens ayant eu au moins un accident)
- `/param = glm` permet un codage des effets comme dans la proc GLM, c'est à dire que la dernière catégorie est prise en référence et mise à zéro.
- `/selection = stepwise` : le modèle va se construire avec la méthode pas à pas d'entrée des effets ; le seuil pour entrer est 0.05 et celui pour rester est de 0.10. (valeurs par défaut de SAS)
- `ctable pprob = (0 to 1 by 0.025)` : SAS imprime la table de classement pour tous les niveaux de probabilité calculés de 0 à 1 par taquets de 0.025.
- `outroc = roc` : SAS crée une table qui sera utilisée pour représenter la courbe ROC.

```
/* LOGISTIC SAS v8 */
proc logistic data = in.assur descending;

class cusag sexe clang age3m cpost2m bm2m_1 puis2m dpoli2m dcons2m /param=glm;
/* param = glm pour avoir un codage des effets comme la proc GLM */

model sinist = cusag sexe clang age3m cpost2m bm2m_1 puis2m dpoli2m dcons2m
/selection = stepwise ctable pprob = (0 to 1 by 0.025) outroc = roc;

format sinist sinistf. cusag cusagf. age agef. sexe sexef. clang clangf.
cpost cpostf. bm bmf. bm_1 bm_1f. dpolic dpolicf. typev typevf.
puis puisf. vcatalg vcatalgf. dcons dconsf. gart2 gart2f. gart3 gart3f.
prsin8m prsin8mf. chsin8m chsin8mf. age8m age8mf.
bm9m bm9mf. bm9m_1 bm9m_1f. puis12m puis12mf. age3m age3mf. cpos2mt cpost2mf.
bm2m_1 bm2m_1f. dpoli2m dpoli2mf. puis2m puis2mf. dcons2m dcons2mf. prim2m prim2mf;

run;

goptions cback=white colors=(black)border;

axis1 length=4in; axis2 order=(0 to 1 by .1) length=3.5in; plot _SENSIT_ * _1MSPEC_ /haxis=axis1
vaxis=axis2;run; proc gplot data = roc;

proc gplot data = roc;symbol1 i=join v=none;
plot _SENSIT_ * _1MSPEC_ /haxis=axis1 vaxis=axis2;
run;quit;
```

Les résultats (extraits)  
The LOGISTIC Procedure

Model Information

Data Set	IN.ASSUR	
Response Variable	SINIST	Sinistralité RC - SNB11 2-3
Number of Response Levels	2	
Number of Observations	1106	
Link Function	Logit	
Optimization Technique	Fisher's scoring	

Response Profile

Ordered Value	SINIST	Total Frequency
1	> 1 sin	550
2	0 sin	556

NOTE: 1 observation was deleted due to missing values for the response or explanatory variables.

Stepwise Selection Procedure

Class Level Information

Class	Value	Design Variables		
		1	2	3
CUSAG	Privé	1	0	
	Profess.	0	1	
SEXE	féminin	1	0	0
	masculin	0	1	0
	sexe autre (Soci	0	0	1
CLANG	Lang franç.	1	0	
	Lang néerland.	0	1	
AGE3M	1890-1949	1	0	0
	1950-73	0	1	0
	Naiss ???	0	0	1
CPOST2M	1	1	0	
	2	0	1	
BM2M_1	Autres B-M (-1)	1	0	
	B-M 1 (-1)	0	1	
PUIS2M	10-39 Puis	1	0	
	40-349 Puis	0	1	
DPOLI2M	<86 Police	1	0	
	autres polices	0	1	
DCONS2M	33-89 DCOS	1	0	
	90-91 DCOS	0	1	



### Summary of Stepwise Selection

Step	Effect		DF	Number		Score	Wald	
	Entered	Removed		In	Chi-Square	Chi-Square	Pr > ChiSq	
1	BM2M_1		1	1	495.2161	.	<.0001	
2	DCONS2M		1	2	62.0832	.	<.0001	
3	AGE3M		2	3	50.3087	.	<.0001	
4	CPOST2M		1	4	21.4766	.	<.0001	
5	CUSAG		1	5	16.2115	.	<.0001	
6	PUIS2M		1	6	8.8740	.	0.0029	
7	DPOLI2M		1	7	6.2583	.	0.0124	

### Summary of Stepwise Selection

Step	Variable Label
1	Bonus-malus Année -1 (2 mod) - GBM1
2	Année de construction du véhicule (2 mod)
3	Age de l assuré (3 mod) - DNAI 8-9
4	Code postal souscripteur (2 mod) - POSS2
5	Code usage - CUSA 5-6
6	Puissance du véhicule (2 mod) - PUIS 32-
7	Date effet Police (2 mod) - DPEP 26-27

### Type III Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
CUSAG	1	11.4673	0.0007
AGE3M	2	44.0929	<.0001
CPOST2M	1	22.0972	<.0001
BM2M_1	1	165.0321	<.0001
PUIS2M	1	8.7026	0.0032
DPOLI2M	1	6.2087	0.0127
DCONS2M	1	41.7697	<.0001

## Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Chi-Square	Pr > ChiSq
Intercept		1	0.0120	0.3204	0.0014	0.9700
CUSAG	Privé	1	-0.8367	0.2471	11.4673	0.0007
<b>CUSAG</b>	<b>Profess.</b>	<b>0</b>	<b>0</b>	<b>.</b>	<b>.</b>	<b>.</b>
AGE3M	1890-1949	1	0.3211	0.2270	2.0001	0.1573
<b>AGE3M</b>	<b>1950-73</b>	<b>1</b>	<b>1.5757</b>	<b>0.2484</b>	<b>40.2329</b>	<b>&lt;.0001</b>
AGE3M	Naiss ???	0	0	.	.	.
CPOST2M	1	1	0.9010	0.1917	22.0972	<.0001
CPOST2M	2	0	0	.	.	.
<b>BM2M_1</b>	<b>Autres B-M (-1)</b>	<b>1</b>	<b>2.4587</b>	<b>0.1914</b>	<b>165.0321</b>	<b>&lt;.0001</b>
BM2M_1	B-M 1 (-1)	0	0	.	.	.
PUIS2M	10-39 Puis	1	-0.7347	0.2490	8.7026	0.0032
PUIS2M	40-349 Puis	0	0	.	.	.
DPOLI2M	<86 Police	1	-0.5021	0.2015	6.2087	0.0127
<b>DPOLI2M</b>	<b>autres polices</b>	<b>0</b>	<b>0</b>	<b>.</b>	<b>.</b>	<b>.</b>
DCONS2M	33-89 DCOS	1	-1.3476	0.2085	41.7697	<.0001
<b>DCONS2M</b>	<b>90-91 DCOS</b>	<b>0</b>	<b>0</b>	<b>.</b>	<b>.</b>	<b>.</b>

### Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
CUSAG Privé vs Profess.	0.433	0.267	0.703
AGE3M 1890-1949 vs Naiss ???	1.379	0.883	2.151
AGE3M 1950-73 vs Naiss ???	4.834	2.971	7.866
CPOST2M 1 vs 2	2.462	1.691	3.585
BM2M_1 Autres B-M (-1) vs B-M 1 (-1)	11.689	8.033	17.010
PUIS2M 10-39 Puis vs 40-349 Puis	0.480	0.294	0.781
DPOLI2M <86 Police vs autres polices	0.605	0.408	0.898
DCONS2M 33-89 DCOS vs 90-91 DCOS	0.260	0.173	0.391

### Association of Predicted Probabilities and Observed Responses

Percent Concordant	90.3	Somers' D	0.817
Percent Discordant	8.6	Gamma	0.827
Percent Tied	1.2	Tau-a	0.409
Pairs	305800	c	0.908

### La table de classement

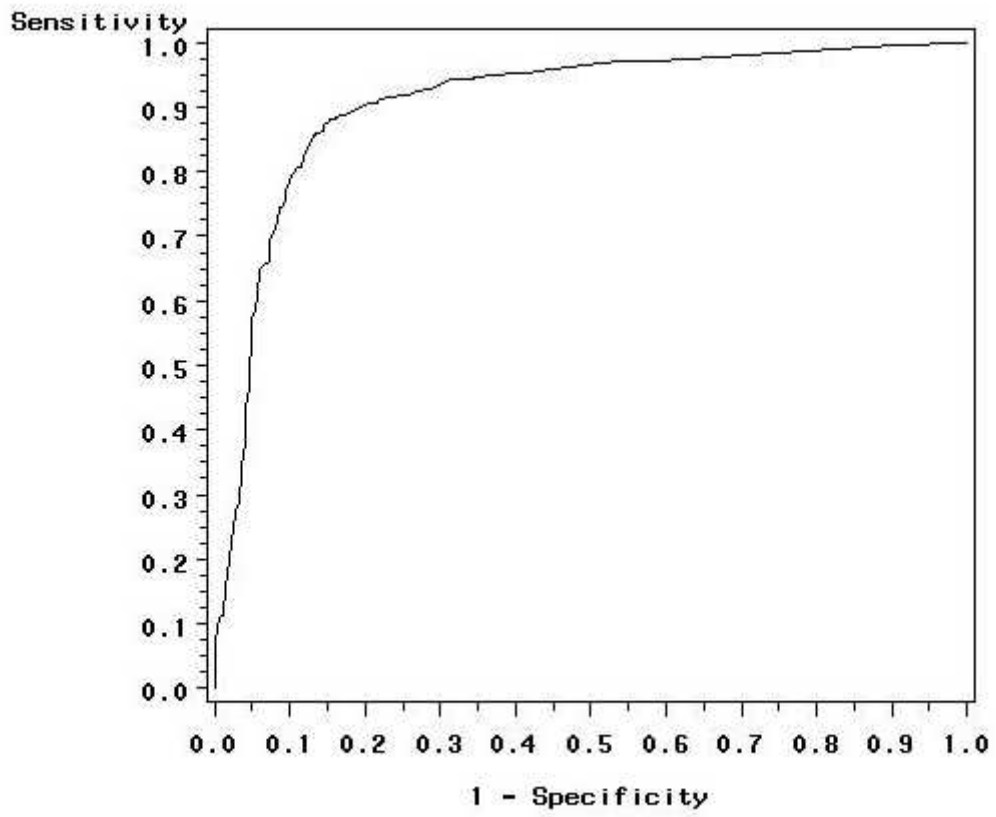
- La spécificité est ici la proportion de personnes ayant eu un sinistre ou plus dans l'année correctement affectées au groupe des '> 1 sinistre'
- La sensibilité est ici la proportion de personnes n'ayant pas eu de sinistre dans l'année et correctement affectées au groupe des 'sinistres'

Ces pourcentages sont calculés pour toutes les valeurs de Y du modèle (Prob level). Par exemple, on se fixe la règle de décision suivante : affecter un individu au groupe des '>1 sinistre' si la probabilité calculée d'appartenance à ce groupe est supérieure à 0.5, et l'affecter au groupe des '0 sinistre' sinon.

Les performances de cette règle dans les deux groupes se lisent dans cette table (en souligné) : on classe correctement 87.6% des '> 1 sinistre' et 84.7% des '0 sinistres'.

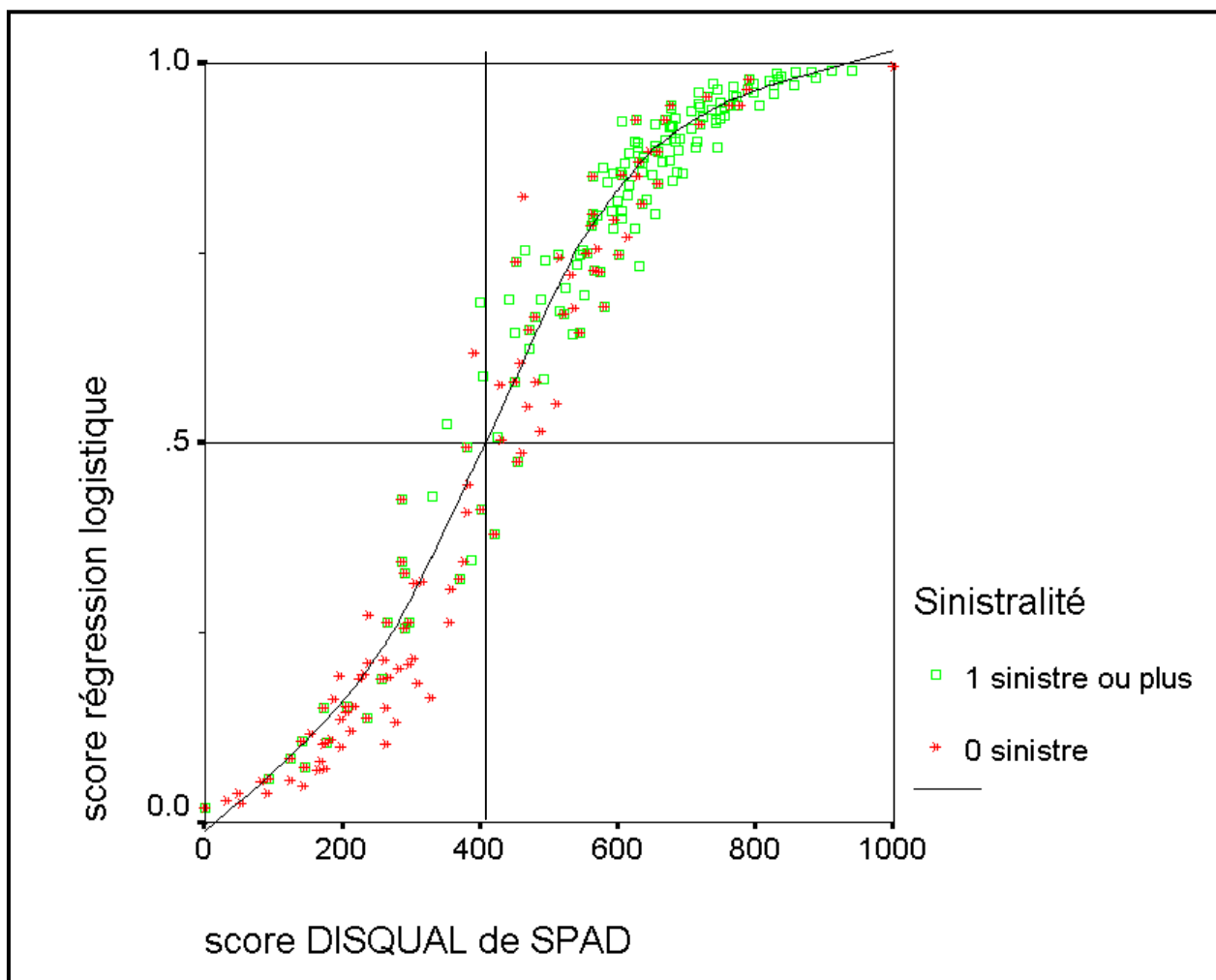
Prob Level	Correct		Incorrect		Percentages				
	Event	Non-Event	Event	Non-Event	Correct	Sensitivity	Specificity	False POS	False NEG
0.000	550	0	556	0	49.7	100.0	0.0	50.3	.
0.025	550	0	556	0	49.7	100.0	0.0	50.3	.
0.050	546	72	484	4	55.9	99.3	12.9	47.0	5.3
0.075	535	244	312	15	70.4	97.3	43.9	36.8	5.8
0.100	528	295	261	22	74.4	96.0	53.1	33.1	6.9
0.125	527	307	249	23	75.4	95.8	55.2	32.1	7.0
0.150	519	363	193	31	79.7	94.4	65.3	27.1	7.9
0.175	519	366	190	31	80.0	94.4	65.8	26.8	7.8
0.200	518	375	181	32	80.7	94.2	67.4	25.9	7.9
0.225	511	396	160	39	82.0	92.9	71.2	23.8	9.0
0.250	509	397	159	41	81.9	92.5	71.4	23.8	9.4
0.275	502	413	143	48	82.7	91.3	74.3	22.2	10.4
0.300	502	431	125	48	84.4	91.3	77.5	19.9	10.0
0.325	501	433	123	49	84.4	91.1	77.9	19.7	10.2
0.350	498	433	123	52	84.2	90.5	77.9	19.8	10.7
0.375	493	441	115	57	84.4	89.6	79.3	18.9	11.4
0.400	491	452	104	59	85.3	89.3	81.3	17.5	11.5
0.425	491	454	102	59	85.4	89.3	81.7	17.2	11.5
0.450	485	454	102	65	84.9	88.2	81.7	17.4	12.5
0.475	485	466	90	65	86.0	88.2	83.8	15.7	12.2
<b>0.500</b>	<b>482</b>	<b>471</b>	<b>85</b>	<b>68</b>	<b>86.2</b>	<b>87.6</b>	<b>84.7</b>	<b>15.0</b>	<b>12.6</b>
0.525	479	473	83	71	86.1	87.1	85.1	14.8	13.1
0.550	479	475	81	71	86.3	87.1	85.4	14.5	13.0
0.575	472	475	81	78	85.6	85.8	85.4	14.6	14.1
0.600	471	481	75	79	86.1	85.6	86.5	13.7	14.1
0.625	471	483	73	79	86.3	85.6	86.9	13.4	14.1
0.650	444	483	73	106	83.8	80.7	86.9	14.1	18.0
0.675	423	500	56	127	83.5	76.9	89.9	11.7	20.3
0.700	420	503	53	130	83.5	76.4	90.5	11.2	20.5
0.725	417	503	53	133	83.2	75.8	90.5	11.3	20.9
0.750	410	507	49	140	82.9	74.5	91.2	10.7	21.6
0.775	375	515	41	175	80.5	68.2	92.6	9.9	25.4
0.800	364	516	40	186	79.6	66.2	92.8	9.9	26.5
0.825	319	523	33	231	76.1	58.0	94.1	9.4	30.6
0.850	312	529	27	238	76.0	56.7	95.1	8.0	31.0
0.875	256	531	25	294	71.2	46.5	95.5	8.9	35.6
0.900	181	536	20	369	64.8	32.9	96.4	10.0	40.8
0.925	136	540	16	414	61.1	24.7	97.1	10.5	43.4
0.950	62	550	6	488	55.3	11.3	98.9	8.8	47.0
0.975	20	555	1	530	52.0	3.6	99.8	4.8	48.8
1.000	0	556	0	550	50.3	0.0	100.0	.	49.7

*La courbe ROC*

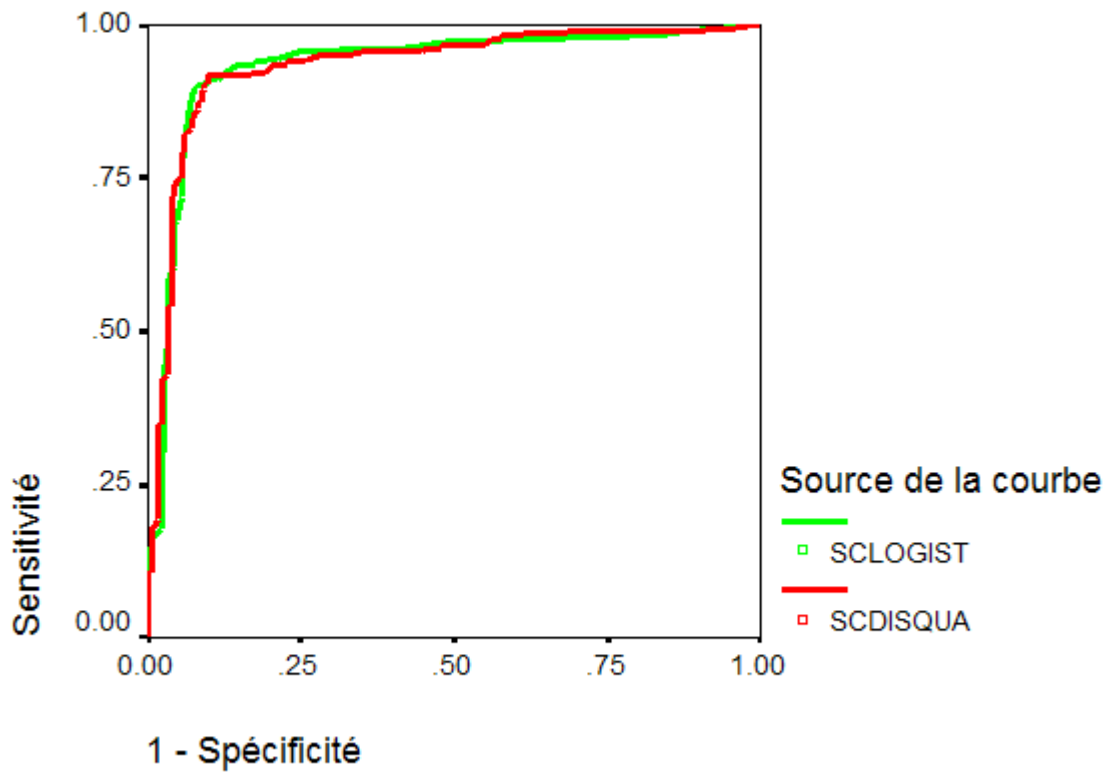


## I.8 Croisement des deux scores

Croisement des deux fonctions de score (DISQUAL et LOGISTIC) pour les 741 individus de l'échantillon d'apprentissage



## Courbe ROC



Zone sous la courbe

Variable(s) de	Zone
SCDISQUA	.934
SCLOGIST	.933