

EXEMPLE : FAILLITE D'ENTREPRISES

Cet exemple a pour objectif d'étudier la faillite d'entreprises. Les données proviennent de l'ouvrage de R.A.Johnson et D.W Wichern : "Applied Multivariate Statistical Analysis », Prentice-Hall, 1982"

Ces données financières annuelles ont été recueillies sur des entreprises approximativement deux ans avant leur faillite, et à peu près à la même époque, sur des sociétés financièrement solides. On dispose de quatre ratios pour décrire la situation financière de 46 entreprises.

X1 = Cash flow/ Dette totale

X2 = Revenu net / Total des actifs au bilan

X3 = Actif réalisable et disponible / Passif courant

X4 = Actif réalisable et disponible / Ventes nettes

Y= 1 si faillite, 2 sinon

Dans ce document nous présentons les résultats obtenus avec la procédure LOGISTIC du logiciel SAS.

Tableau 1 : les données

Obs.	X1	X2	X3	X4	Y	num
1	-0.4485	-0.4106	1.0865	0.4526	Faillite	1-
2	-0.5633	-0.3114	1.5134	0.1642	Faillite	2-
3	0.0643	0.0156	1.0077	0.3978	Faillite	3-
4	-0.0721	-0.0930	1.4544	0.2589	Faillite	4-
5	-0.1002	-0.0917	1.5644	0.6683	Faillite	5-
6	-0.1421	-0.0651	0.7066	0.2794	Faillite	6-
7	0.0351	0.0147	1.5046	0.7080	Faillite	7-
8	-0.0653	-0.0566	1.3737	0.4032	Faillite	8-
9	0.0724	-0.0076	1.3723	0.3361	Faillite	9-
10	-0.1353	-0.1433	1.4196	0.4347	Faillite	10-
11	-0.2298	-0.2961	0.3310	0.1824	Faillite	11-
12	0.0713	0.0205	1.3124	0.2497	Faillite	12-
13	0.0109	0.0011	2.1495	0.6969	Faillite	13-
14	-0.2777	-0.2316	1.1918	0.6601	Faillite	14-
15	0.1454	0.0500	1.8762	0.2723	Faillite	15-
16	0.3703	0.1098	1.9941	0.3828	Faillite	16-
17	-0.0757	-0.0821	1.5077	0.4215	Faillite	17-
18	0.0451	0.0263	1.6756	0.9494	Faillite	18-
19	0.0115	-0.0032	1.2602	0.6038	Faillite	19-
20	0.1227	0.1055	1.1434	0.1655	Faillite	20-
21	-0.2843	-0.2703	1.2722	0.5128	Faillite	21-
22	0.5135	0.1001	2.4871	0.5368	OK	22+
23	0.0769	0.0195	2.0069	0.5304	OK	23+
24	0.3776	0.1075	3.2651	0.3548	OK	24+

Obs.	X1	X2	X3	X4	Y	num
25	0.1933	0.0473	2.2506	0.3309	OK	25+
26	0.3248	0.0718	4.2401	0.6279	OK	26+
27	0.3132	0.0511	4.4500	0.6852	OK	27+
28	0.1184	0.0499	2.5210	0.6925	OK	28+
29	-0.0173	0.0233	2.0538	0.3484	OK	29+
30	0.2169	0.0779	2.3489	0.3970	OK	30+
31	0.1703	0.0695	1.7973	0.5174	OK	31+
32	0.1460	0.0518	2.1692	0.5500	OK	32+
33	-0.0985	-0.0123	2.5029	0.5778	OK	33+
34	0.1398	-0.0312	0.4611	0.2643	OK	34+
35	0.1379	0.0728	2.6123	0.5151	OK	35+
36	0.1486	0.0564	2.2347	0.5563	OK	36+
37	0.1633	0.0486	2.3080	0.1978	OK	37+
38	0.2907	0.0597	1.8381	0.3786	OK	38+
39	0.5383	0.1064	2.3293	0.4835	OK	39+
40	-0.3330	-0.0854	3.0124	0.4730	OK	40+
41	0.4785	0.0910	1.2444	0.1847	OK	41+
42	0.5603	0.1112	4.2918	0.4443	OK	42+
43	0.2029	0.0792	1.9936	0.3018	OK	43+
44	0.4746	0.1380	2.9166	0.4487	OK	44+
45	0.1661	0.0351	2.4527	0.1370	OK	45+
46	0.5808	0.0371	5.0594	0.1268	OK	46+

**QUESTION 1 : CONSTRUCTION D'UN MODELE DE REGRESSION LOGISTIQUE
RELIANT Y AUX VARIABLES X1, X2, X3, X4.**

Procédure LOGISTIC

Model Information

Table	WORK.DONNEES
Variable de réponse	Y
Nombre de niveaux de réponse	2
Modèle	logit binaire
Technique d'optimisation	Score de Fisher

Nombre d'observations lues 46

Nombre d'observations utili 46

Response Profile

Ordered Value	Y	Total Frequency
1	Faillite	21
2	OK	25

La probabilité modélisée est Y='Faillite'.

Le tableau 2 (obtenu à l'aide de l'instruction simple) permet de comparer le groupe 1 et le groupe 2 sur l'ensemble des variables explicatives. On remarque que pour le groupe 2 les moyennes des variables X1, X2, X3 sont plus élevées. En revanche les moyennes de X4 sont très voisines pour les deux groupes.

TABLEAU 2 : Descriptive Statistics for Continuous Variables

Variable	Y	Mean	Standard Deviation	Minimum	Maximum
X1	Faillite	-0.068824	0.210329	-0.563300	0.370300
	OK	0.235356	0.217601	-0.333000	0.580800
	Total	0.096491	0.261499	-0.563300	0.580800
X2	Faillite	-0.081862	0.144622	-0.410600	0.109800
	OK	0.055052	0.048099	-0.085400	0.138000
	Total	-0.007452	0.123628	-0.410600	0.138000
X3	Faillite	1.367490	0.405645	0.331000	2.149500
	OK	2.593892	1.022724	0.461100	5.059400
	Total	2.034013	1.006203	0.331000	5.059400
X4	Faillite	0.438114	0.210756	0.164200	0.949400
	OK	0.426440	0.161804	0.126800	0.692500
	Total	0.431770	0.183682	0.126800	0.949400

TABLEAU 3 : Model Fit Statistics

Criterion	Intercept Only	Constante et Covariables
AIC	65.421	37.457
SC	67.250	46.601
-2 Log	63.421	27.457

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DDL	Pr > Khi-2
Rapport de vrais	35.9639	4	<.0001
Score	22.3037	4	0.0002
Wald	12.2118	4	0.0158

TABLEAU 4: Analysis of Maximum Likelihood Estimates

Parameter	DDL	Estimate	Standard Error	Wald Chi-Square	Pr > Khi-2
Intercept	1	5.3711	2.3928	5.0385	0.0248
X1	1	-7.1975	6.0172	1.4308	0.2316
X2	1	3.8096	13.7273	0.0770	0.7814
X3	1	-3.4285	1.2151	7.9610	0.0048
X4	1	2.9266	3.0803	0.9027	0.3421

Les tests du rapport des vraisemblances, du score ou de Wald présentés dans le tableau 3 conduisent tous au rejet de l'hypothèse H0 de nullité de l'ensemble des coefficients. $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$. Néanmoins (TABLEAU 4) seule la variable X3 a un apport marginal significatif dans le modèle complet.

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
X1	<0.001	<0.001	99.094
X2	45.132	<0.001	>999.999
X3	0.032	0.003	0.351
X4	18.663	0.045	>999.999

TABLEAU 5: Association of Predicted Probabilities and Observed Responses

Pourcentage concordant	94.1	D de Somers	0.884
Pourcentage discordant	5.7	Gamma	0.885
Pourcentage lié	0.2	Tau-a	0.448
Paires	525	c	0.942

QUESTION 2 : Sélection ascendante des variables

Étape 0. Intercept saisi :

$$-2 \text{ Log L} = 63.421$$

Analysis of Maximum Likelihood Estimates

Parameter	DDL	Estimate	Standard Error	Wald Chi-Square	Pr > Khi-2
Intercept	1	-0.1744	0.2960	0.3469	0.5558

Residual Chi-Square Test

Chi-Square	DDL	Pr > Khi-2
22.3037	4	0.0002

Analysis of Effects Eligible for Entry

Effect	DDL	Score Chi-Square	Pr > Khi-2
X1	1	15.7859	<.0001
X2	1	14.3089	0.0002
X3	1	17.3317	<.0001
X4	1	0.0471	0.8281

Etape 1. Effet X3 saisi :

Model Fit Statistics

Criterion	Intercept Only	Constante et Covariables
AIC	65.421	39.341
SC	67.250	42.998
-2 Log L	63.421	35.341

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DDL	Pr > Khi-2
Rapport de vrais	28.0804	1	<.0001
Score	17.3317	1	<.0001
Wald	11.7393	1	0.0006

Analysis of Maximum Likelihood Estimates

Parameter	DDL	Estimate	Standard Error	Wald Chi-Square	Pr > Khi-2
Intercept	1	6.0691	1.8131	11.2047	0.0008
X3	1	-3.3815	0.9869	11.7393	0.0006

