

Exemple 3.2 : Modélisation MCO incluant des variables explicatives dichotomiques (reprise de l'étude du jugement des enseignants)

resm= jugement que porte l'enseignant sur la valeur scolaire de ses élèves en mathématiques
scom= scores obtenus par les élèves aux épreuves nationales de CE2 en mathématiques
(performances scolaires effectives de début d'année)

Syntaxe détaillée:

```
proc reg Indique la procédure de régression
data=work.jugement Spécifie le nom du tableau de données
model resm=scom Sert à définir le modèle (variable dépendante=régresseurs)
run; Permet de lancer les estimations
```

Syntaxe SAS:

```
proc reg data=work.jugement;
model resm=scom;
run;
```

Sorties SAS :

Procédure REG					
Modèle : MODEL1					
Variable dépendante : RESM					
Nombre d'observations lues		342			
Nombre d'observations utilisées		342			
Analyse de variance					
Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F
Modèle	1	619.67209	619.67209	267.58	<.0001
Erreur	340	787.38054	2.31583		
Total sommes corrigées	341	1407.05263			
Valeurs estimées des paramètres					
Root MSE		1.52178	R carré	0.4404	
Moyenne dépendante		6.61404	R car. ajust.	0.4388	
Coeff Var		23.00840			
Valeurs estimées des paramètres					
Variable	DDL	Valeur estimée des paramètres	Erreur type	Valeur du test t	Pr > t
Intercept	1	1.82980	0.30383	6.02	<.0001
SCOM	1	0.11614	0.00710	16.36	<.0001

Lignes 22 et 23 : Cet exemple porte sur 342 élèves

Ligne 36 : Indique le coefficient de détermination (R^2) non ajusté

Ligne 37 : Indique le coefficient de détermination (R^2) ajusté

Ligne 44 : la colonne $Pr > |t|$ indique la probabilité critique associée à l'impact très significatif du score de début de CE2 en mathématiques sur le jugement de l'enseignant.

Introduction dans le modèle d'une autre variable (régresseur) explicative :

retard= élèves en retard scolaire

Syntaxe SAS:

```
proc reg data=work.jugement;
model resm=scom retard;
run;
```

Sorties SAS :

Procédure REG					
Modèle : MODEL1					
Variable dépendante : RESM					
Nombre d'observations lues		342			
Nombre d'observations utilisées		342			
Analyse de variance					
Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F
Modèle	2	660.21398	330.10699	149.84	<.0001
Erreur	339	746.83865	2.20306		
Total sommes corrigées	341	1407.05263			
Root MSE		1.48427	R carré	0.4692	
Moyenne dépendante		6.61404	R car. ajust.	0.4661	
Coeff Var		22.44125			
Valeurs estimées des paramètres					
Variable	DDL	Valeur estimée des paramètres	Erreur type	Valeur du test t	Pr > t
Intercept	1	2.25802	0.31270	7.22	<.0001
SCOM	1	0.10938	0.00710	15.40	<.0001
retard	1	-0.98343	0.22925	-4.29	<.0001

Ligne 80 : Indique le coefficient de détermination (R^2) non ajusté

Ligne 81 : Indique le coefficient de détermination (R^2) ajusté

Ligne 92: On relève un impact significatif et négatif du retard scolaire sur le jugement de l'enseignant.

Introduction dans le modèle d'une autre variable (régresseur) explicative :

psup= origine sociale des élèves (favorisée ou non). Enfants dont le père est cadre supérieur/profession libérale

Syntaxe SAS:

```

103 proc reg data=work.jugement;
104 model resm=scom retard psup;
105 run;

```

106 Sorties SAS :

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

Procédure REG

Modèle : MODEL1

Variable dépendante : RESM

Nombre d'observations lues342

Nombre d'observations utilisées342

Analyse de variance

Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F
Modèle	3	675.38691	225.12897	104.00	<.0001
Erreur	338	731.66573	2.16469		
Total sommes corrigées	341	1407.05263			

Root MSE	1.47129	R carré	0.4800
Moyenne dépendante	6.61404	R car. ajust.	0.4754
Coeff Var	22.24495		

Valeurs estimées des paramètres

Variable	DDL	Valeur estimée des paramètres	Erreur type	Valeur du test t	Pr > t
Intercept	1	2.27116	0.31000	7.33	<.0001
SCOM	1	0.10620	0.00714	14.87	<.0001
retard	1	-0.90615	0.22911	-3.96	<.0001
psup	1	0.54902	0.20737	2.65	0.0085

139 *Ligne 125:* Indique le coefficient de détermination (R^2) non ajusté

140 *Ligne 126 :* Indique le coefficient de détermination (R^2) ajusté

141 *Ligne 138 :* Indique que, à score et déroulement de la scolarité donnés, les enfants dont
142 l'origine sociale est favorisée sont mieux jugés par leur enseignant que les autres.