

Jean-Michel Marin

Un exemple de sortie SAS

La sortie SAS ci-dessous donne le résultat d'une régression linéaire multiple où la variable à expliquer est le loyer charges comprises d'appartements de la région parisienne et les deux variables explicatives sont la surface et le nombre de pièces.

Le programme permettant de l'obtenir est le suivant :

```
proc reg data=sasuser.logemt2;
model loyercc= surface nbp / selection = rsquare cp covb dw ss1 ss2;
plot r.*p.;
plot(r. student.)*obs.;
plot(u95. l95.)*surface='- ' p.*surface/overlay symbol='*';
run;
```

Les options suivantes permettent d'afficher :

- covb : $\widehat{\sigma^2} (X X')^{-1}$;
- ss1, ss2 : le type de somme de carrés (par défaut ss2) ;
- selection = rsquare : implique le calcul de tous les 2^k modèles avec k régresseurs ; l'estimation et les tests sont uniquement réalisés pour le modèle ayant le C_p de Mallows le plus élevé si l'option cp est spécifiée, le rsquare le plus élevé sinon ; (cette option n'est pas conseillée si le nombre de régresseur est élevé, on utilisera plutôt selection = forward ou selection = backward dans ce cas) ;
- dw : la valeur de la statistique de Durbin-Watson et le graphe des résidus.

La première commande plot assure l'affichage du graphe des résidus contre les valeurs estimées La deuxième commande plot assure l'affichage de deux graphiques l'un avec les résidus l'autre avec les résidus studentisés contre le numéro de l'observation. La dernière commande plot assure en affichage '*' les valeurs estimées et en affichage '-' les bornes de l'intervalle de confiance le tout contre la variable surface.

The SAS System

N = 141 Regression Models for Dependent Variable: LOYERCC			
Number in Model	R-square	C(p)	Variabes in Model
1	0.78303565	2.25953	SURFACE
1	0.58504432	129.34115	NBP

2 0.78499798 3.00000 SURFACE NBP



Model: MODEL1
 Dependent Variable: LOYERCC

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	2	432597260.07	216298630.04	251.927	0.0001
Error	138	118483469.67	858575.86719		
C Total	140	551080729.74			
Root MSE	926.59369	R-square	0.7850		
Dep Mean	4489.95745	Adj R-sq	0.7819		
C.V.	20.63703				

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > T
INTERCEP	1	1132.550468	176.58758190	6.414	0.0001
SURFACE	1	105.872351	9.34543669	11.329	0.0001
NBP	1	-220.551653	196.51949263	-1.122	0.2637

Variable	DF	Type I SS	Type II SS
INTERCEP	1	2842520220	35316161
SURFACE	1	431515856	110190612
NBP	1	1081404	1081404

Covariance of Estimates

COVB	INTERCEP	SURFACE	NBP
INTERCEP	31183.174082	-284.5442708	-8770.668013
SURFACE	-284.5442708	87.337186889	-1629.857154
NBP	-8770.668013	-1629.857154	38619.910983

Durbin-Watson D 1.668
 (For Number of Obs.) 141
 1st Order Autocorrelation 0.149



