

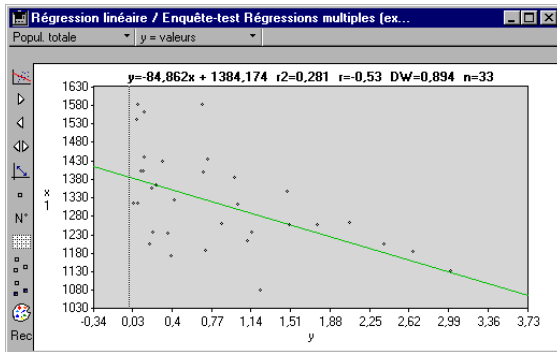
Modalisa 4

Création, analyse de questionnaires et d'entretiens pour MAC & PC



**Corrélations,
Régressions simples
et multiples**

CORRELATION ET REGRESSION LINEAIRE SIMPLE



Modifications et ajouts depuis la version 3.7 :

Statistique de Durbin-Watson : $DW = 2*(1-r)$ pour $r =$ coefficient d'autocorrélation (Tomassone & aliud, p.125).

Possibilité de ne pas calculer et de ne pas afficher la droite de régression (= nuage de points seulement).

Données affichées sur l'axe des abscisses : valeurs de y (par défaut), résidus ou résidus réduits (si régression calculée).

Choix de sous-population.

Passage à la variable (x) suivante et précédente.

Choix de variable (x) sur liste.

Transposition : x devient y

Création de recodage (si régression calculée) : Valeurs calculées de y , résidus, résidus réduits

Représentation des individus : Points, numéros, densité par zone, double-clic pour avoir la liste des individus présents dans une zone et visualiser leurs réponses.

Marquage de sous-population.

REGRESSION MULTIPLE

Choisissez l'article « Régression multiple » dans le menu « Analyse ».

Choisissez d'abord la variable à expliquer (y).

Puis sélectionnez les variables explicatives (régresseurs ou x).

Les tableaux de résultats

Page principale

Log(y)		Régresseurs (x)		Correl./y
Equation de la régression: $\text{Log}(y) = 11,003 + (-0,004 * x1) + (-0,054 * x2) + (0,068 * x3) + (-1,291 * x4) + (0,232 * x5) + (-0,353 * x6) + (-0,243 * x7) + (0,179 * x8) + (-1,282 * x9) + (-0,423 * x10)$				
n = 33				
	Coeff	E.T.	t	p
1. x1	-0,00444	0,002	2,853 ***	p=0,005
2. x2	-0,05402	0,022	2,468 **	p=0,011
3. x3	0,06777	0,099	0,682	p=0,255
4. x4	-1,29140	0,562	2,298 **	p=0,015
5. x5	0,23228	0,104	2,225 **	p=0,018
6. x6	-0,35311	1,565	0,226	p=0,409
7. x7	-0,24308	1,004	0,242	p=0,403
8. x8	0,17993	0,236	0,756	p=0,232
9. x9	-1,28171	0,865	1,482 *	p=0,075
10. x10	-0,42372	0,734	0,586	p=0,285

$R^2 =$ coefficient de détermination, $F =$ statistique de Fisher, $p =$ probabilité associée au F .

Statistique de Durbin-Watson : $DW = 2*(1-r)$ pour $r =$ coefficient d'autocorrélation (Tomassone & aliud, p.125).

Equation de la droite de régression multiple.

Individus concernés, non réponse et sous-population le cas échéant.

Liste des régresseurs et de leur coefficient de corrélation avec y .

Tableau d'analyse de la variance

$SC_{reg} =$ somme des carrés de la régression, $mSC_{reg} =$ moyenne des carrés de la régression

$SC_{res} =$ somme des carrés des résidus, $mSC_{res} =$ moyenne des carrés des résidus.

$S_{tot} =$ somme des carrés de la variable expliquée (y).

Liste des coefficients

Coefficients, écart-type des coefficients, valeur du t de Student, probabilité associée au t .

Si la probabilité associée au t est inférieure au égale à 0,1 la colonne « t » comprendra une ou plusieurs étoiles.

*** : $p \leq 0,01$

- : $0,01 < p \leq 0,05$
- : $0,05 < p \leq 0,1$

Inflation R2

Bouton « Inflation R2 ». En cliquant sur ce bouton vous demandez le calcul de la colonne correspondante.

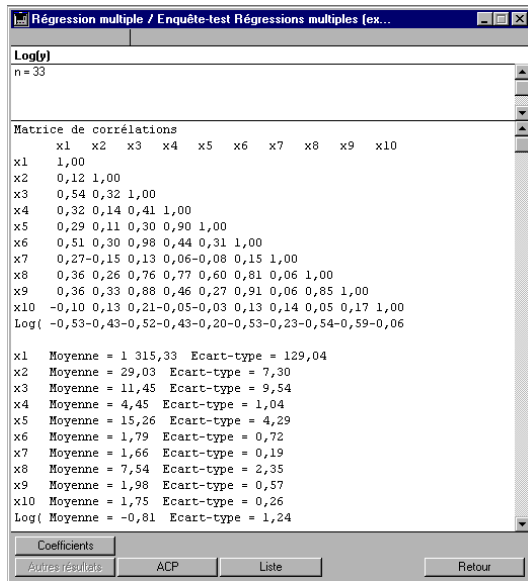
Les coefficients de détermination (R2) sont calculés pour chaque x en fonction des p-1 autres régresseurs.

La colonne « Inflation R2 » contient le facteur d'inflation : $f_i = 1/(1-R2_i)$.

Si tel régresseur est très lié aux autres, R2 sera proche de 1.

« Nous pouvons donner comme règle grossière que tout facteur d'inflation supérieur à 10 (...) peut entraîner certaines imprécisions dans les estimations, imprécisions dues à une collinéarité trop importante entre variables explicatives. » (Tomassone & aliud, p.150).

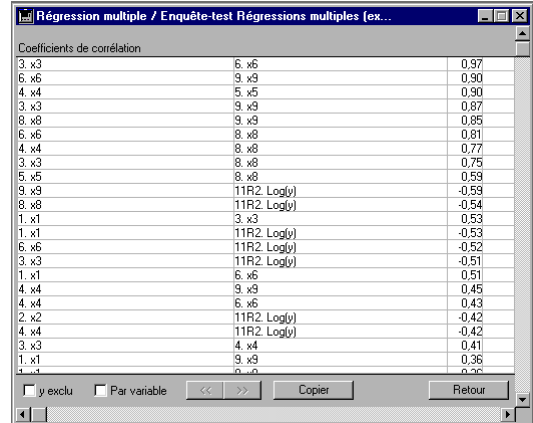
Matrice de corrélation, moyennes et écart-types



Bouton « Corrélation ».

Corrélations en liste

Bouton « Corrélation » puis « Liste » : Liste des corrélations triées par ordre décroissant (valeur absolue) Bouton « Par variable » : affiche toutes les corrélations d'une variable.



Autres tableaux de résultats

Le bouton « Autre résultats » affiche d'autres tableaux spécifiques à telle ou telle méthode.

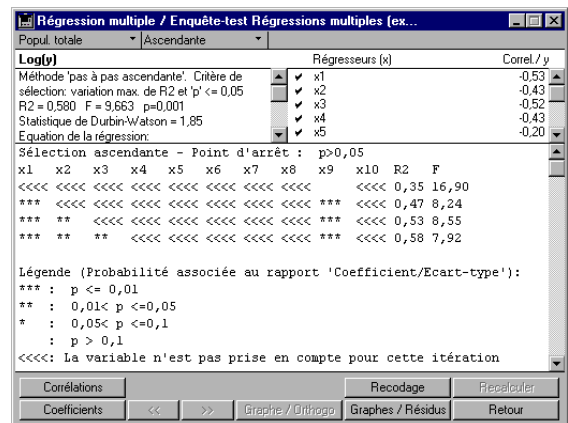
Ces tableaux sont détaillés ci-dessous dans le paragraphe « Méthodes ». Pour passer d'un tableau à l'autre utilisez les flèches.

LES METHODES

Standard

Le vecteur des coefficients est donné par la formule : $(X'X)^{-1} X'y$

Pas à pas ascendante



Choix de p (0 < p <= 1)

Algorithme :

Pas 1 : sélection du régresseur dont le coefficient de corrélation (valeur absolue) avec y est le plus élevé.

Ajout du régresseur restant dont le 't' est le plus élevé lors du calcul incluant les régresseurs déjà introduits.

Arrêt : la valeur de 'p' pour le prochain régresseur à introduire est supérieure au 'p' choisi.

Le bouton « Autre résultat » affiche un tableau présentant une ligne par itération. Pour chaque régresseur la probabilité associée au rapport 'Coefficient/Ecart-type' est indiquée avec les symboles suivants :

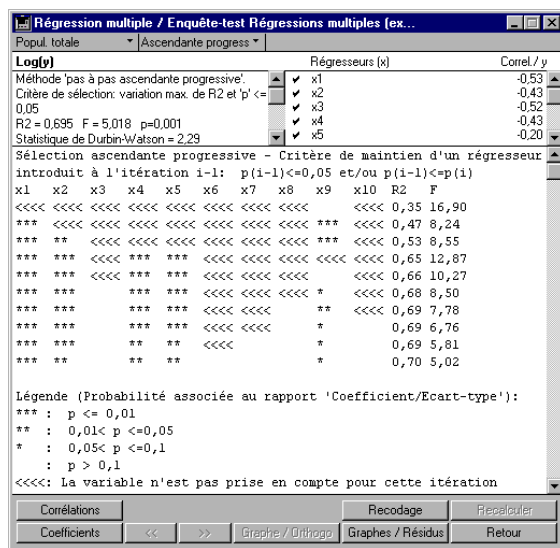
- : $p \leq 0,01$
- : $0,01 < p \leq 0,05$
- : $0,05 < p \leq 0,1$
- : $p > 0,1$

<<<<: La variable n'est pas prise en compte pour cette itération

Deux colonnes supplémentaires affichent le R2 et le F de l'itération correspondante.

Pas à pas ascendante progressive

Choix de p ($0 < p \leq 1$)



Algorithme :

Pas 1 : sélection du régresseur dont le coefficient de corrélation (valeur absolue) avec y est le plus élevé.

Pas suivants : Ajout du régresseur restant dont le 't' est le plus élevé lors du calcul incluant les régresseurs déjà introduits.

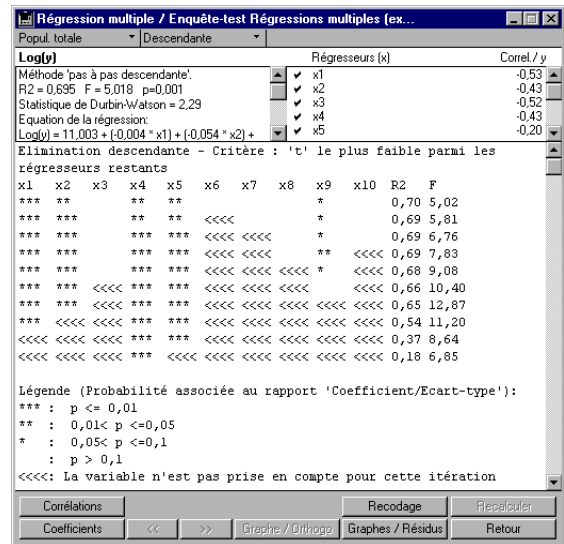
Retour sur les régresseurs déjà inclus : lors de l'introduction d'un nouveau régresseur, un x déjà pris en compte peut être est replacé dans le groupe des régresseurs à calculer si au moins une des deux conditions suivante est remplie :

- La valeur de « t » est inférieure à celle du dernier;

- La probabilité associée au « t » est supérieure au « p » choisi.

Le bouton « Autres résultats » affiche un tableau présentant une ligne par itération (cf Pas à pas ascendante).

Pas à pas descendante

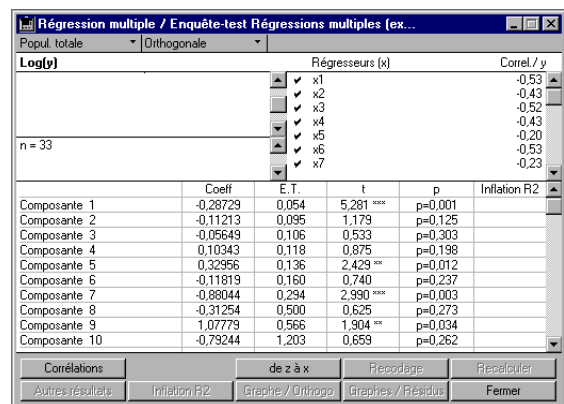


A chaque pas : suppression du régresseur dont le « t » est le plus faible lors du calcul de la régression avec les x restants en lice.

Le bouton « Autre résultat » affiche un tableau présentant un ligne par itération (cf Pas à pas ascendante).

Orthogonalisée

Les variables x sont transformées par une analyse en composantes principales. La régression est calculée sur ces nouvelles variables z.



La première page présente les coefficients correspondant à chaque composante. Les valeurs de t permettent d'identifier les composantes significatives (Tomassone & aliud, p.146-147).

	Coef	E.T.	t	p	Inflation R2
1. x1	-0,46020	0,161	2,852 ***	p=0,005	
2. x2	-0,31694	0,128	2,468 **	p=0,011	
3. x3	0,51915	0,760	0,683	p=0,254	
4. x4	-1,07961	0,471	2,293 **	p=0,015	
5. x5	0,80119	0,361	2,221 **	p=0,018	
6. x6	-0,20347	0,901	0,226	p=0,409	
7. x7	-0,03654	0,151	0,242	p=0,403	
8. x8	0,33782	0,447	0,755	p=0,232	
9. x9	-0,58322	0,393	1,482 *	p=0,075	
10. x10	-0,08848	0,151	0,586	p=0,285	

Le bouton « z vers x » permet de revenir aux régresseurs de départ (cf. Tomassone & aliud, p.148-149).

Bouton « Autres résultats » :

Plusieurs tableaux et graphiques permettent d'interpréter les résultats. Pour passer d'un tableau à l'autre, utilisez les flèches.

c	10	9	8	7	6	5	4	3	2	1
b1	-0,46	-0,44	-0,48	-0,49	-0,21	-0,28	-0,14	-0,10	-0,08	-0,08
b2	-0,32	-0,31	-0,32	-0,32	-0,33	-0,29	-0,05	-0,03	-0,06	-0,05
b3	0,52	0,04	0,37	0,32	-0,11	-0,12	-0,16	-0,15	-0,15	-0,12
b4	-1,08	-0,93	-0,17	-0,19	-0,06	-0,04	-0,02	-0,04	-0,04	-0,10
b5	0,80	0,71	0,22	0,32	-0,02	-0,04	0,01	-0,01	-0,01	-0,08
b6	-0,20	0,38	0,26	0,23	-0,11	-0,11	-0,16	-0,15	-0,15	-0,12
b7	-0,04	-0,08	-0,15	-0,13	-0,21	-0,14	-0,06	-0,09	-0,05	-0,02
b8	0,34	0,20	-0,26	-0,43	-0,06	-0,04	-0,10	-0,11	-0,11	-0,12
b9	-0,58	-0,64	-0,63	-0,40	-0,07	-0,05	-0,14	-0,14	-0,14	-0,12
b10	-0,09	-0,04	-0,04	-0,06	0,05	0,00	0,03	-0,05	-0,06	-0,02

Tableau 1 : Coefficients en fonction du nombre de composantes incluses (cf. Tomassone & aliud, p.152-154).

Pour obtenir le graphique correspondant : bouton « Graphe/Ortho ».

cp	1	2	3	4	5	6	7	8	9	10
cp1	0,8	0,5	0,1	0,2	0,2	0,1	0,0	0,3	0,3	0,0
cp2	0,1	1,2	0,1	1,0	2,3	0,1	2,5	0,1	0,3	5,8
cp3	7,6	19,4	0,0	0,0	0,0	0,0	24,0	0,0	0,1	1,6
cp4	6,2	1,4	0,0	0,3	0,7	0,0	4,9	0,0	0,0	38,8
cp5	12,6	61,7	0,0	0,0	0,3	0,1	5,9	0,4	0,9	0,6
cp6	35,1	14,7	0,1	0,1	0,4	0,0	30,4	0,6	0,3	17,5
cp7	33,2	0,0	3,6	0,9	<<<<<	1,6	2,7	7,8	7,9	5,7
cp8	0,0	0,2	1,1	0,8	20,3	0,4	4,8	39,8	84,7	3,4
cp9	1,8	0,2	5,2	71,9	51,9	0,5	6,7	28,8	0,0	0,0
cp10	2,6	0,6	89,8	24,6	13,8	97,4	17,9	22,3	5,3	26,5
Var.	1,9	1,2	41,7	16,0	9,4	58,6	1,6	14,4	11,2	1,6

Tableau 2 : Proportion de la variance des coefficients de régression (en %).

Ce tableau permet notamment de détecter des collinéarités (Tomassone & aliud, p.151-152).

b1	b2	b3	b4	b5	b6	b7	b8	b9	b10	
b1	1,00									
b2	-0,02	1,00								
b3	-0,27	-0,09	1,00							
b4	0,00	0,02	-0,29	1,00						
b5	-0,12	-0,02	0,20	-0,85	1,00					
b6	0,10	0,07	-0,92	0,42	-0,31	1,00				
b7	-0,29	0,15	0,33	-0,43	0,47	-0,40	1,00			
b8	0,10	0,00	0,34	-0,60	0,19	-0,42	0,18	1,00		
b9	0,04	-0,09	0,08	-0,16	0,38	-0,31	0,25	-0,39	1,00	
b10	0,32	-0,05	-0,51	0,29	-0,30	0,48	-0,39	-0,09	-0,25	1,00

Tableau 3 : Coefficients de corrélation des coefficients de régression (Tomassone & aliud, p.155).

Pseudo-orthogonalisée

Même mode de calcul que pour la régression orthogonalisée.

Ajout d'une valeur k (0<k<=1) sur la diagonale de la matrice X'X.

Bouton « Autres résultats » :

k	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09	0,1	0,25	0,5	0,75	1
b1	-0,46	-0,44	-0,43	-0,42	-0,41	-0,39	-0,39	-0,38	-0,37	-0,36	-0,35	-0,28	-0,22	-0,19	-0,17
b2	-0,32	-0,31	-0,31	-0,31	-0,30	-0,30	-0,29	-0,29	-0,29	-0,29	-0,29	-0,25	-0,21	-0,18	-0,15
b3	0,52	0,30	0,24	0,21	0,18	0,17	0,15	0,14	0,13	0,12	0,11	0,02	-0,02	-0,04	-0,05
b4	-1,08	-0,85	-0,72	-0,63	-0,57	-0,52	-0,47	-0,44	-0,41	-0,39	-0,37	-0,22	-0,14	-0,11	-0,10
b5	0,80	0,66	0,57	0,51	0,46	0,42	0,39	0,36	0,34	0,32	0,30	0,16	0,08	0,05	0,02
b6	-0,20	0,05	0,11	0,13	0,13	0,13	0,12	0,11	0,11	0,10	0,09	0,02	-0,02	-0,04	-0,05
b7	-0,04	-0,07	-0,09	-0,10	-0,10	-0,11	-0,11	-0,12	-0,12	-0,12	-0,12	-0,11	-0,09	-0,09	-0,08
b8	0,34	0,17	0,09	0,04	0,00	-0,02	-0,04	-0,06	-0,07	-0,08	-0,08	-0,11	-0,10	-0,10	-0,09
b9	-0,58	-0,56	-0,52	-0,49	-0,46	-0,44	-0,42	-0,40	-0,38	-0,37	-0,35	-0,24	-0,18	-0,15	-0,13
b10	-0,09	-0,06	-0,05	-0,04	-0,04	-0,03	-0,03	-0,03	-0,03	-0,02	-0,02	-0,02	-0,01	0,00	0,00

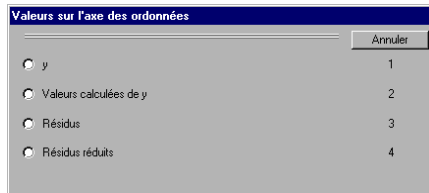
Tableau 1 : Coefficients en fonction des valeurs de k. Pour obtenir le graphique correspondant : bouton « Graphe/Ortho ».

Ce tableau et le graphique permettent d'éliminer les variables dont les coefficients sont faibles ou instables et tendant vers zéro (Tomassone & aliud, p.155-159).

Les tableaux 2 et 3 sont les mêmes que pour la régression orthogonalisée.

Graphiques

Valeurs calculées de y, résidus, résidus réduits

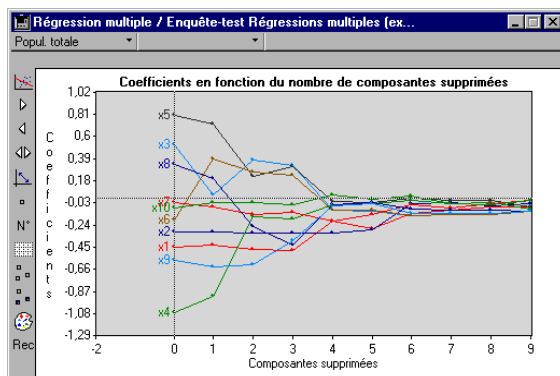


Bouton « Graphe/résidus »

Les options de la régression simple sont disponibles sur le graphique. Restrictions : il n'est pas possible ici de créer un recodage, ni de transposer les variables. Le calcul de la droite de régression n'est pas possible si les résidus sont affichés.

Orthogonalisée

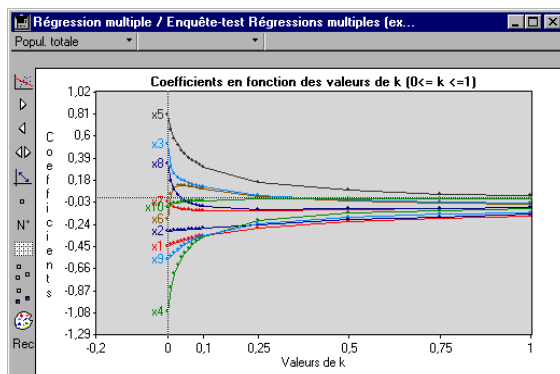
Bouton « Graphe/orthogo » sur la méthode orthogonalisée après avoir transformé les z en x.



Chaque ligne brisée représente les valeurs d'un coefficient par suppression progressive d'une composante.

Pseudo-orthogonalisée

Bouton « Graphe/orthogo » sur la méthode pseudo-orthogonalisée. L'axe des x correspond aux valeurs de k ($0 \leq k \leq 1$). Variation d'un centième entre 0 et 0.1. Puis 0.25, 0.5, 0.75 et 1.



ACP

Bouton « Corrélation » puis « ACP ».

Affiche le mapping d'analyse factorielle en composantes principales. Particulièrement utile pour évaluer la position des régresseurs sur les facteurs significatifs révélés par la méthode orthogonalisée.

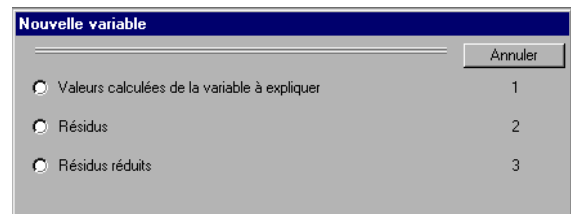
Divers

Sous-sélection de régresseurs

Dans le tableau des régresseurs une coche indique que cette variable est incluse dans l'analyse. En cliquant sur la ligne d'un régresseur vous l'excluez ou l'introduisez selon l'état antérieur. Lorsque votre nouvelle sélection est terminée, cliquez sur le bouton « Recalculer ».

Choix de sous-population

Création de recodage



Valeurs calculées de y, résidus ou résidus réduits.

NB. ce type de recodage ne sera pas mis à jour automatiquement lors de la modification des données brutes.

Impression, exportation et copier-coller

Tous les résultats peuvent être imprimés, exportés dans un fichier ou copiés vers le presse-papier. Pour copier vers le presse-papier utilisez la combinaison de touches « Commande c » (Mac) ou « Contrôle c » (Windows).