



**Département des Sciences de l'Éducation**

Section française

Route de Faucigny 2

CH-1700 Fribourg

# Petit aide-mémoire *SPSS*

*SPSS 6*



# Contenu

<b>Codage des données</b>	<b>5</b>
Présentation des données	5
Passage d'un fichier de <i>Excel</i> à <i>SPSS</i>	5
Utilisation de quelques fonctions sur <i>SPSS</i>	6
Mettre des étiquettes (labels) à une variable et à ses modalités	6
Calculer une nouvelle variable sur la base de données	6
Recoder une variable	7
Séparer un fichier en plusieurs parties	7
Visualisation des résultats	8
Copie des résultats dans <i>Word</i>	8
Enregistrement	8
<b>Quelques fonctions avec <i>SPSS</i> ... et leurs résultats</b>	<b>9</b>
Fonctions descriptives	9
Chi carré	10
Test du <i>t</i> de Student	13
Variables indépendantes	13
Variables appariées	15
ANOVA (plusieurs VD et 1 VI)	17
ANOVA (1 VD et plusieurs VI)	19
Corrélation	22
Régression multiple	24
Conclusion	27



# Codage des données

## Présentation des données

Une ligne par sujet

Une colonne par variable

Tous les résultats en chiffres (pas de lettres !)

Nom de la variable : maximum 8 caractères (commencer par une lettre)

Pour les variables manquantes, le plus simple est de laisser la case vide.

Exemple :

Sexe	Section	Score1	Score2	etc
1	1	10	8	
1	1	8	5	
2	2	5	4	
1	2	7	3	
1	3	2	3	
2	3	5	3	

1 = garçon

2 = fille

1 = pratique

2 = générale

3 = pré-gymnasiale

A éviter (en gras) :

		Scores	
Sexe	Section de l'élève	1	2
<b>g</b>	1	10	8
<b>g</b>	1	8	5
<b>f</b>	2	5	4
<b>g</b>	2	7	3
<b>g</b>	3	2	3
<b>f</b>	3	5	3

## Passage d'un fichier de Excel à SPSS

Il est conseillé d'entrer les données sur une feuille *Excel*, puis de les transférer dans *SPSS*. Pour se faire :

1. Préparez la feuille qui ne contiendra que les données brutes (pas de graphiques, de moyennes par colonne, ...)
2. Toutes les données à traiter doivent se trouver sur la même feuille (ne pas séparer par exemple en créant une feuille pour les filles et une pour les garçons).
3. Enregistrez la feuille en format "texte, séparateur tabulation".

Chemin :

Fichier → Enregistrer sous... → Type du fichier : → Texte (séparateur: tabulation).

[Pour SPSS 8 ou 9, enregistrez votre feuille en format Excel 4].

4. Fermez la feuille (sinon, vous ne pourrez pas l'ouvrir dans SPSS).  
Attention ! En général Lorsque vous fermez la feuille, un messages apparaît. Lisez-le ! On vous informe que le fichier que vous fermez n'est pas en format *Excel 95* ou *98*, et on vous propose de le mettre sous ce format. Cliquez "non", car vous désirez que le format reste en "texte, séparateur tabulation" !
5. Dans SPSS, ouvrez vos données.  
Chemin :  
Fichier → ouvrir → ... (ou cliquez sur l'icône correspondante)

### Utilisation de quelques fonctions sur SPSS

Le logiciel SPSS travaille sur plusieurs feuilles :

- une feuille de données (data) : cette feuille contient les données sur lesquelles vous effectuez des traitements statistiques (on ne peut en avoir qu'une seule ouverte à la fois)
- une(des) feuille(s) de sortie (output) : contient le résultat des tests effectués.

D'autres feuilles (syntaxe, graphique, ...) peuvent être créées. A ce sujet, nous vous recommandons d'utiliser le logiciel *Excel* pour réaliser vos graphiques car il permet de faire des modifications pour une mise en page pertinente dans *Word*.

### **Fonction : Mettre des étiquettes (labels) à une variable et à ses modalités**

Chemin : (Sélectionnez la colonne) Data → Define Variable... → Labels...

Variable Label : vous pouvez donner à votre variable un nom plus grand que les 8 caractères en en-tête de colonne.

Value : Entrez la modalité de la variable (exemple : 1)

Value Label : Entrez le nom de la modalité (exemple : garçon)

Cliquez sur Add et répétez les deux dernières commandes pour chaque modalité, puis cliquez sur Continue et OK.

En cliquant à présent sur l'avant-dernière icône à droite (représentant une petite étiquette), vous pouvez visualiser soit la valeur, soit le nom associé à cette valeur. Quelque soit la configuration sur l'écran, dans les feuilles de résultats, vous verrez ces noms, forts utiles pour se repérer dans l'amas de chiffres !

### **Fonction : Calculer une nouvelle variable sur la base de données**

Chemin : Transform → Compute...

Target Variable : entrez le nom de la nouvelle variable que vous calculez

Numeric Expression : vous pouvez ici combiner les variables existantes avec différents opérateurs mathématiques.

Cliquez ensuite sur OK : le logiciel vous crée la nouvelle variable (et vous la place dans la dernière colonne)

Exemple :

Target Variable

Somme

=

Numeric Expression

Score1 + Score2

Exemple :

Pour passer d'une variable codée négativement à une variable retournée (positive).

Question : "Je ne suis pas motivé par le travail scolaire"

Réponse : 1 = Pas du tout d'accord à 7 = Tout à fait d'accord

L'élève qui répond 7 est le moins motivé.

Si l'on veut retourner cette variable afin que l'élève le plus motivé obtienne le score maximal selon le tableau suivant :

Codage effectué par l'élève	1	2	3	4	5	6	7
Score "réel"	7	6	5	4	3	2	1

Target Variable

Score

=

Numeric Expression

8 - codage

### Fonction : Recoder une variable

Exemple : on veut passer d'un score allant de 1 à 10 à un score de 1 à 3 suivant le tableau suivant :

Anciennes modalités	1	2	3	4	5	6	7	8	9	10
Nouvelles modalités	1		2				3			

Chemin : Transform → Recode → Into Different Variables...

Input Variable : sélectionnez la variable à recoder et faites passer la variable dans la boîte "Input variable" à l'aide de la flèche noire

Output Variable : insérer le nom (8 caractères max) et le label complet (pas obligatoire) puis cliquez sur Change

Old and New Values...

Entrez l'ancienne valeur (case de gauche)

Entrez la nouvelle valeur (case de droite)

Cliquez sur Add et répétez les deux dernières commandes (y compris le "Add") pour toutes les modalités (y compris celles qui ne changent pas)

Cliquez sur Continue, puis sur OK.

### Fonction : Séparer un fichier en plusieurs parties

On utilise cette fonction pour effectuer par exemple des tests avec les garçons et les filles séparément.

Chemin : Data → Split File...

Cliquez sur "Repeat analysis for each group"

Sélectionnez la variable (par exemple "Sexe") et faites la passer à droite à l'aide de la flèche noire

Cliquez sur OK.

A partir de ce moment, tous les tests effectués se feront à double (pour cet exemple), soit une fois pour les garçons et une fois pour les filles.  
Pour retourner à une analyse sur l'ensemble des données :

Chemin :

Data → Split File...  
Cliquez sur "Analyze all cases"  
Cliquez sur OK.

### **Visualisation des résultats**

Chemin :

Window  
Vous pouvez soit voir les résultats, soit les données.

### **Copie des résultats dans *Word***

Pour éviter d'imprimer de nombreuses pages à partir de *SPSS*, vous pouvez sélectionner les résultats, puis les copier dans le presse-papier.

Chemin :

Edit → Copy  
Allez dans *Word* pour coller ce qui a été copié dans *SPSS*.  
Pour avoir une bonne présentation (si les données viennent de *SPSS 6.1*, car il n'y a plus ce problème avec *SPSS 8.1*), mettez les dans la police "courrier", taille 9.

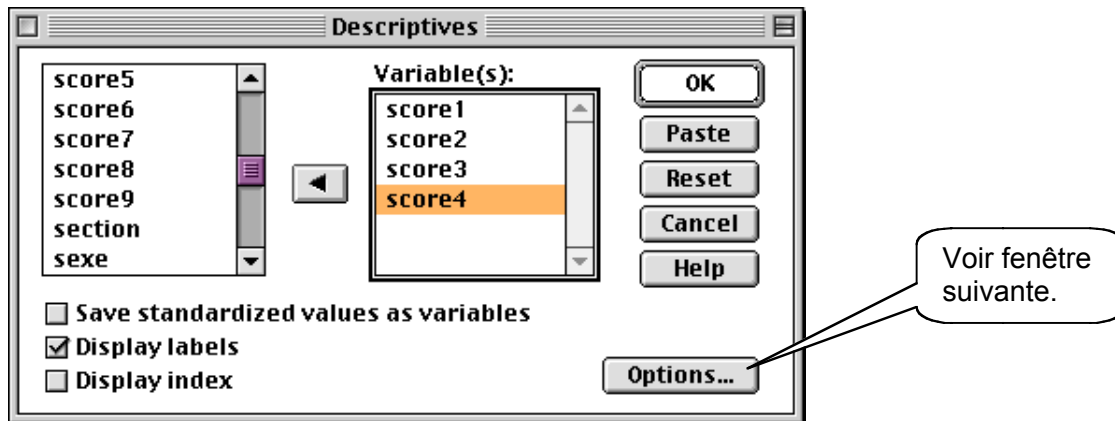
### **Enregistrement**

Attention, il faut enregistrer séparément la feuille de donnée et la(les) feuille(s) des résultats !

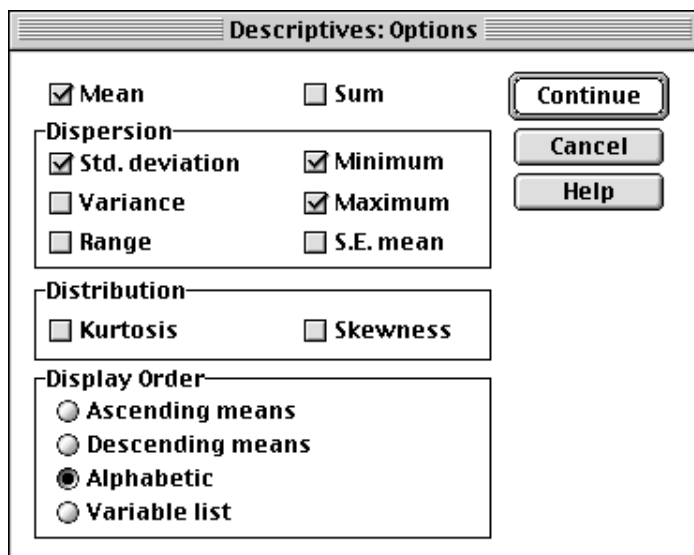
# Quelques fonctions avec SPSS ... et leurs résultats

## Fonctions descriptives

Chemin : Statistics → Summarize → Descriptives...



Sélectionner dans la fenêtre de gauche les variables, puis les passer à droite en utilisant la flèche (entre les deux fenêtres).



On peut cocher ici les statistiques descriptives qui nous intéressent ...

Mean : moyenne  
Sum : somme (en colonne)  
Std. deviation : écart-type  
Range : rang  
Minimum : chiffre minimum  
Maximum : chiffre maximum  
S.E. mean : erreur-type

... les indices de distribution des résultats (Kurtosis et Skewness) ...

... et l'ordre de présentation des résultats (ici, par ordre alphabétique des noms des variables).

## Résultats

Nombre de cases (sujets) prises en compte.

```
Number of valid observations (listwise) =      250.00
```

Variable	Mean	Std Dev	Minimum	Maximum	Valid N	Label
SCORE1	3.23	.84	1.00	4.00	250	
SCORE2	2.36	.98	1.00	4.00	250	
SCORE3	2.02	.85	1.00	4.00	250	
SCORE4	2.69	1.09	1.00	4.00	250	

Nombre de cases prises en compte pour chaque variable.

Si l'on souhaite avoir les scores pour les garçons et les filles séparément, il suffit d'utiliser la fonction "Split File" (voir p. 7).

---

## Chi carré ( $\chi^2$ )

### Comparer une distribution observée à une distribution théorique.

Le principe général consiste à analyser l'écart existant entre la distribution théorique postulée et la distribution empirique obtenue<sup>1</sup>.

Les deux variables étudiées sont nominales.

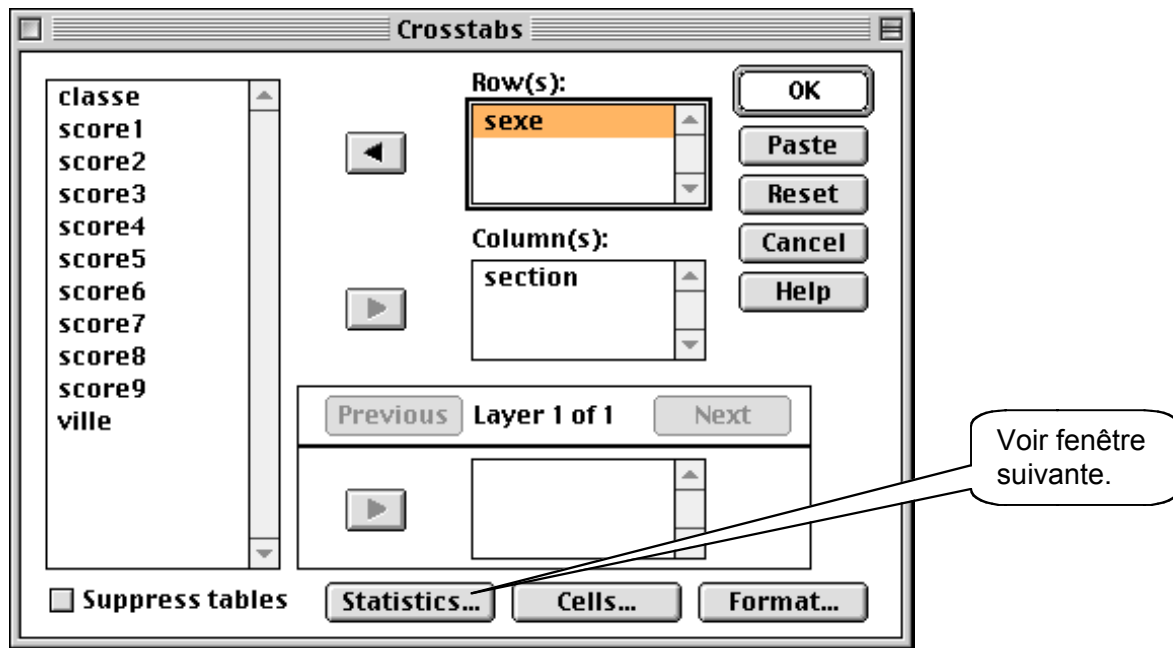
---

<sup>1</sup> Vous pouvez vous référer au cours de statistiques I dans lequel l'explicitation du tableau des fréquences observées (ou distribution empirique) et celui des fréquences attendues (ou distribution théorique) est indispensable pour le calcul "à la main" du Chi2. Attention à ne pas confondre le terme "fréquence" (fréquence absolue, ici) avec les pourcentages (ou fréquences relatives).

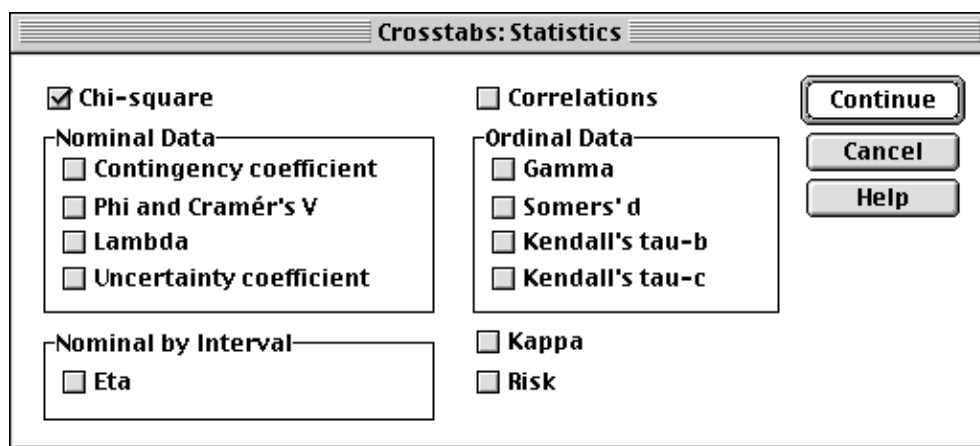
Comme exemple, nous testons l'hypothèse suivante :

Hypothèse : La répartition des garçons et des filles dans les trois sections diffère d'une distribution théorique.

Chemin : Statistics → Summarize → Crosstabs...



On sélectionne les variables à mettre en ligne et en colonne (pour le Chi2, une seule variable en ligne et une seule en colonne). Il faut encore sélectionner le test du Chi2 :



On coche dans la case correspondante (en haut à droite) pour obtenir dans la feuille de résultats la valeur du Chi carré de Pearson<sup>2</sup>.

<sup>2</sup> On différencie le Chi carré de Pearson (Karl Pearson, 1900) d'autres tests donnant eux aussi une statistique  $\chi^2$ .

## Résultats

SEXE by SECTION		SECTION			Page 1 of 1
Count		1.00	2.00	3.00	Row Total
SEXE					
Garçon	1.00	44	45	35	124 49.6
Fille	2.00	50	42	34	126 50.4
Column Total		94	87	69	250 100.0

Chi-Square	Value	DF	Significance
Pearson	.48495	2	.78468
Likelihood Ratio	.48520	2	.78458
Mantel-Haenszel test for linear association	.28697	1	.59217

Minimum Expected Frequency - 34.224

Number of Missing Observations: 0

On obtient donc le tableau croisé (avec le nombre de sujets correspondants à l'intérieur de chaque case).

La valeur du  $\chi^2$  est de 0.485 (arrondi), le degré de liberté est de 2, et la probabilité associée à ce test est de 0.79 donc non significatif (en prenant un seuil de 5%).

### Condition d'utilisation du test :

**Minimum Expected Frequency :** lorsque ce chiffre est inférieur à 5 (valeur correspondant à la plus petite fréquence absolue trouvée dans le tableau de la distribution théorique), le test n'est pas valide (quelque soit la probabilité).

**Number of Missing Observation<sup>3</sup> :** donne le nombre de valeurs manquantes.

<sup>3</sup> On retrouve cette information pour la majorité des statistiques effectuées.

## Test du *t* de Student (comparaison des moyennes deux groupes)

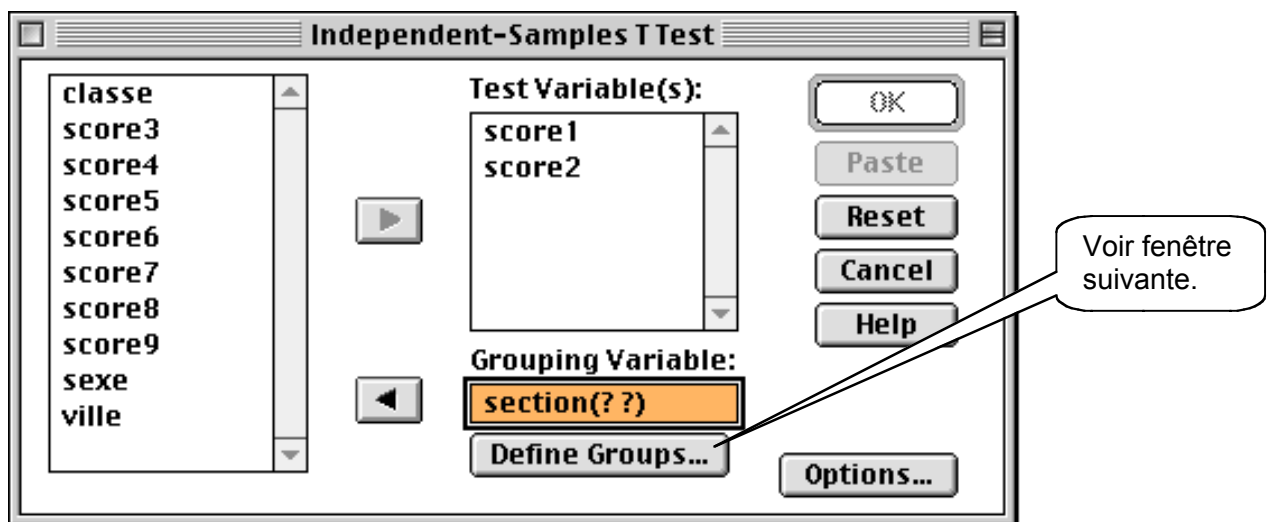
Les variables sont quantitatives.

Si les variables sont indépendantes

Exemple : Comparaison entre le groupe des élèves de pré-gymnastique et ceux de section pratique pour les variables Score1 et Score2.

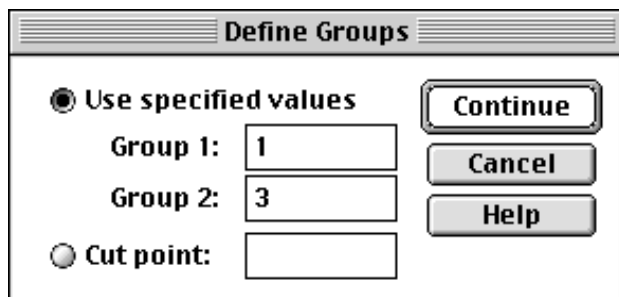
Hypothèse : Les pré-gymnastiques auront pour les deux variables une moyenne significativement inférieure aux élèves de pratique<sup>4</sup>.

Chemin : Statistics → Compare Means → Independent-Samples T Test...



Test Variable(s) : variable(s) dépendante(s) (on peut en mettre plusieurs et un calcul sera effectué pour chaque variable séparément).

Grouping Variable : variable indépendante (définir les deux modalités).



Définir les deux modalités de la variable indépendante sur lesquelles la variable dépendante sera comparée (ici "1" pour pratique et "3" pour pré-gymnastique).

<sup>4</sup> De manière formelle, on devrait mentionner pour cet exemple deux hypothèses (une pour chaque variable séparément).

Lorsque l'on veut comparer entre-elles par exemple trois sections sur la base des moyennes des élèves en fin d'année, il faut répéter l'opération trois fois et comparer tout d'abord les sections 1 et 2, puis 1 et 3, et finalement 2 et 3.

### Résultats (ici, uniquement pour le Score1)

Variable	Number of Cases	Mean	SD	SE of Mean
SCORE1				
Prégymna	94	3.0957	.917	.095
Pratique	69	3.3188	.757	.091

Mean Difference = -.2231

Levene's Test for Equality of Variances: F= 1.606 P= .207

Variances	t-value	df	2-Tail Sig	SE of Diff	95% CI for Diff
Equal	-1.65	161	.080	.135	(-.490, .044)
Unequal	-1.70	158.71	.071	.131	(-.482, .036)

Moyennes

Premier tableau :  
partie descriptive.

Test de  
Student.

Test d'homogénéité  
des variances.

La première chose à regarder : les moyennes des deux groupes pour savoir si les élèves de pratique ont effectivement un score supérieur à ceux de pré-gymnastique (selon l'hypothèse que nous avons posée).

Dans nos résultats, les élèves de pratique obtiennent 3.32, alors que ceux de pré-gymnastique n'ont que 3.10 (arrondi). La différence entre ces deux moyennes est donnée entre les deux tableaux, soit -.022.

Ensuite, le test d'homogénéité des variances permet de sélectionner dans le tableau du bas quelle ligne choisir. Lorsque les variances sont considérées homogènes (comme c'est le cas ici, vu que la probabilité du "Levene's Test for Equality of Variances" est supérieure à 0.05 puisqu'il est de 0.21), on choisit alors la première ligne (celle où les variances sont "Equal"), sinon, on prend la deuxième ligne.

La valeur du  $t$  de Student est de 1.65 (on peut laisser tomber le signe moins), le degré de liberté est de 161 et la probabilité est de 0.08 (ou 8%).

Vu que nous avons posé une hypothèse dite "orientée", c'est-à-dire que nous avons spécifié le sens de la différence (les élèves de pratiques ont un score supérieur), nous pouvons diviser cette probabilité par deux. On trouve alors une probabilité de 4% donc inférieur à notre seuil de 5%. L'hypothèse est donc confirmée.

Exemple d'une hypothèse non orientée (ou bilatérale, ou à deux queues) :

Il y a une différence entre les pré-gymnasiales et les pratiques pour la variable Score1 (mais on n'indique pas quel est le groupe qui est supérieur).

En général, la majorité des hypothèses sont orientées, car le chercheur s'attend à une différence en faveur d'un groupe (par exemple un effet positif d'un cours de remédiation).

Si les variables sont codées de manière ordinale (et non quantitative), on trouve le test correspondant (U de Mann-Whitney)

Chemin : Statistics → Nonparametric Tests → 2 Independent Samples...

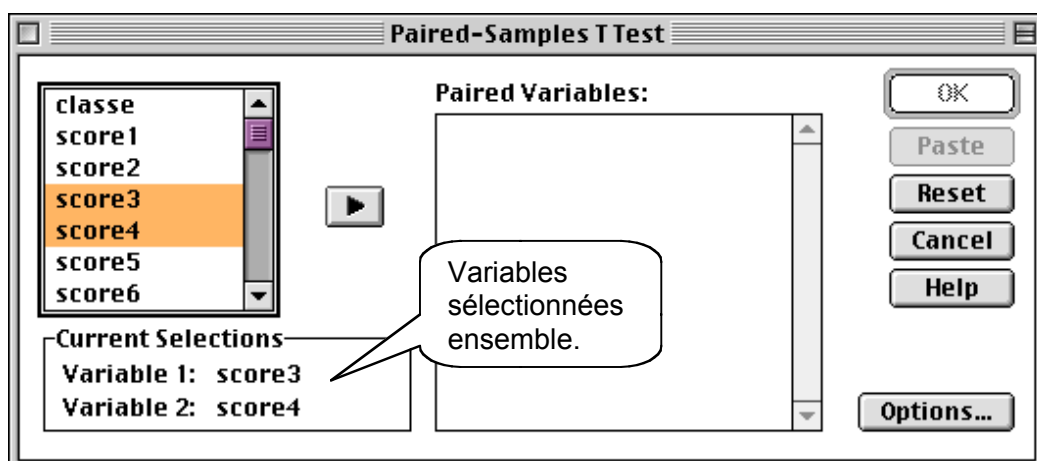
### Si les variables sont appariées (pairees)

Exemple : Comparaison entre le Score3 (par exemple un pré-test) et le Score4 (le post-test correspondant).

Ce sont donc bien les mêmes individus à qui on attribue une valeur au Score 3 et une valeur au Score4 (variables paires).

Hypothèse : Il y a une amélioration significative entre le pré-test et le post-test, autrement dit, le Score4 sera supérieur au Score3 (hypothèse orientée, unilatérale, ou à une queue).

Chemin : Statistics → Compare Means → Paired-Samples T Test...



Pour faire passer dans la fenêtre de droite les deux variables concernées par la comparaison à faire, on doit impérativement les sélectionner ensemble. Pour cela, il faut maintenir enfoncée la touche ⌘ ("pomme") et sélectionner les variables en cliquant.

Les variables sélectionnées se placent dans la fenêtre en bas à gauche (cf. copie de l'écran de la page précédente), il reste ensuite à les faire passer à droite à l'aide de la flèche.

On peut faire exécuter à la suite plusieurs comparaisons deux à deux en sélectionnant dès le départ l'ensemble des comparaisons à effectuer.

## Résultats

t-tests for Paired Samples

Variable	Number of pairs	Corr	2-tail Sig	Mean	SD	SE of Mean
SCORE3	250	-.235	.000	2.0200	.852	.054
SCORE4				2.6880	1.090	.069

Paired Differences			t-value	df	2-tail Sig
Mean	SD	SE of Mean			
-.6680	1.533	.097	-6.89	249	.000
95% CI (-.859, -.477)					

Le Score4 (moyenne de 2.69) a effectivement une moyenne supérieure au Score3 (moyenne de 2.02), et la différence est significative car la probabilité est inférieure à 5% (ici,  $p = 0.000$  ... avec certainement quelques chiffres après, mais le résultat est donné avec trois chiffres après la virgule seulement).

Si les variables sont codées de manière ordinale (et non quantitative), on trouve le test correspondant (Wilcoxon)

Chemin : Statistics → Nonparametric Tests → 2 Related Samples...

## ANOVA (plusieurs VD et 1 VI) (analyse de variance)

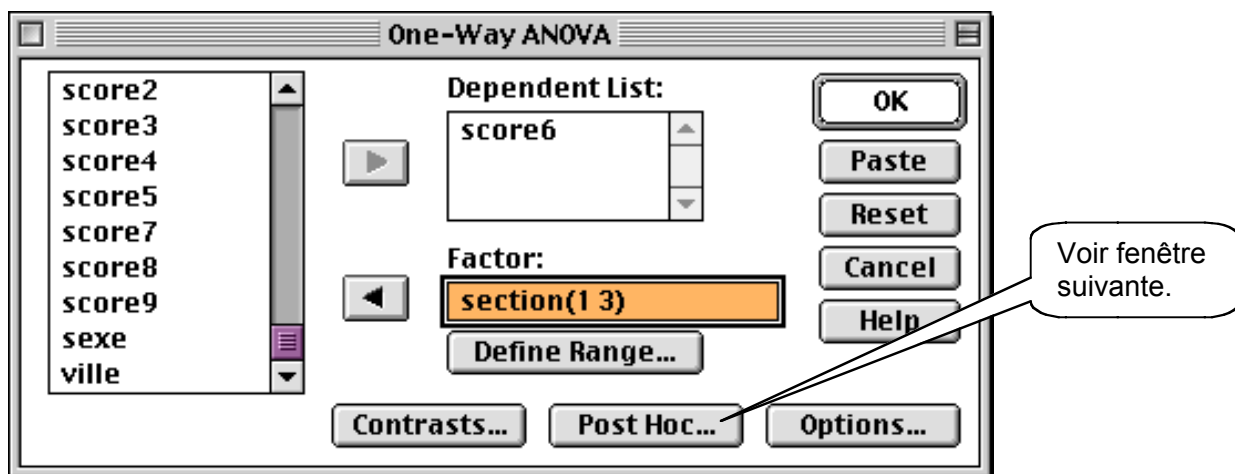
Comparer les scores de plusieurs groupes indépendants de sujets sur une variable quantitative.

De manière (très) simplifiée, on regarde si la variation entre les groupes est plus importante que la variation à l'intérieur des groupes. Si on trouve de grandes différences entre les groupes, mais qu'à l'intérieur de chacun des groupes, les sujets ont approximativement le même score (variation faible), l'analyse de variance (ANOVA) risque d'être significative.

Exemple : Comparaison entre les trois sections pour le score1.

Hypothèse : Il y a au moins une différence significative entre deux des trois groupes sur la variable Score6.

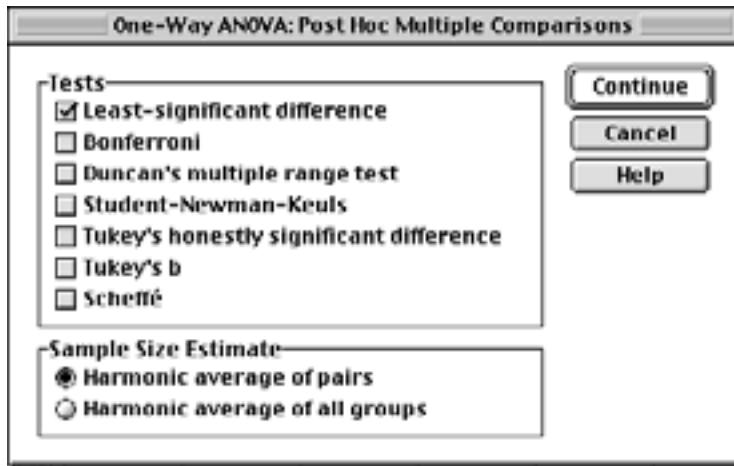
Chemin : Statistics → Compare Means → One-Way ANOVA...



De manière similaire au *t* de Student, on insère la(les) variable(s) dépendante(s) ainsi que le facteur (variable indépendante). Il faut aussi définir sur quelles modalités s'effectuera le test. Ici, il faut donner le chiffre minimum et le chiffre maximum<sup>5</sup>, par exemple, si on effectue une ANOVA entre les sections 1, 2 et 3, on entrera 1 et 3 (case "Define Range").

L'analyse de variance diffère du test de Student par le fait qu'elle peut prendre en compte plusieurs groupes, et aussi tester les différences deux à deux, ce qui évite de répéter le test de Student plusieurs fois.

<sup>5</sup> Il faut bien entendu y penser lors du codage des données !



Les post-hoc permettent des comparaisons deux à deux si (et uniquement si) l'analyse de variance est significative.

On coche donc ici le test "Least-significant difference" qui nous indiquera (dans la feuille des résultats) les paires significativement différentes.

## Résultats

```

      - - - - - O N E W A Y - - - - -

Variable  SCORE6
By Variable  SECTION

Analysis of Variance

Source          D.F.      Sum of Squares      Mean Squares      F Ratio      F Prob.
-----
Between Groups      2          19.0397           9.5198           9.8461      .0001
Within Groups     247         238.8163            .9669
Total              249         257.8560

      - - - - - O N E W A Y - - - - -

Variable  SCORE6
By Variable  SECTION

Multiple Range Tests:  LSD test with significance level .05

The difference between two means is significant if
MEAN(J)-MEAN(I) >= .6953 * RANGE * SQRT(1/N(I) + 1/N(J))
with the following value(s) for RANGE: 2.79

(*) Indicates significant differences which are shown in the lower triangle

          G G G
          r r r
          p p p

          3 2 1

Mean      SECTION
2.6667    Grp 3
2.8506    Grp 2
3.3191    Grp 1      * *
  
```

Degrés de liberté.

F de Fisher.

Probabilité associée au F de Fisher.

Comparaisons deux à deux :  
Groupe 1 différent (significativement) du groupe 2 et du groupe 3.

Le test est significatif (la probabilité de 0.0001 est inférieure au seuil de 5%).

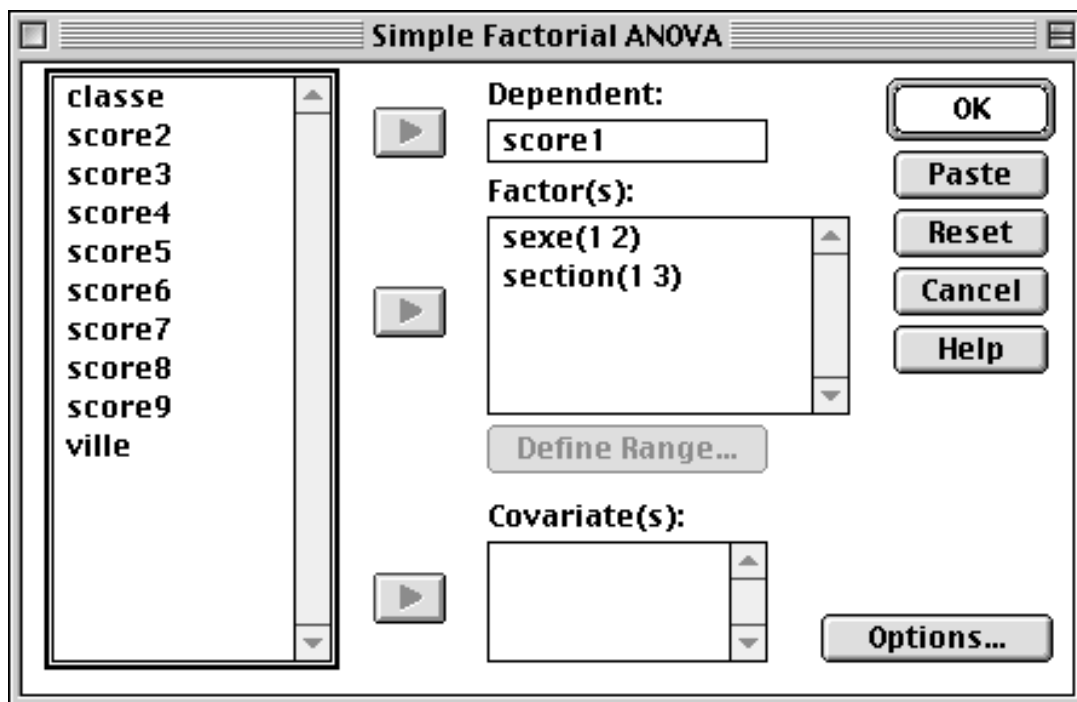
Les post-hoc donnent les paires significativement différentes. Dans le tableau croisé, une étoile est placée à l'intersection des lignes et colonnes dont les variables sont différentes.

Si l'analyse de variance n'est pas significative, ce tableau croisé n'apparaît pas dans la feuille des résultats.

---

## ANOVA (1 VD et plusieurs VI) (analyse de variance)

Chemin : Statistics → ANOVA Models → Simple Factorial...



De manière identique à la "One-Way ANOVA" (pp.17-18), on sélectionne les variables et on définit leurs modalités.

On peut insérer dans ce modèle plusieurs facteurs (ou variables indépendantes), calculer les effets d'interaction, mais les post-hoc (tels qu'ils ont été présentés auparavant) ne sont plus possibles avec ce type d'analyse.

Afin d'avoir deux types de résultats différents, nous présentons successivement les résultats de l'analyse de variance pour le Score1 et le Score4 (variables dépendantes), avec le Sexe et la Section comme variables indépendantes.

## Résultats pour la variable Score1

```


*** ANALYSIS OF VARIANCE ***

      SCORE1
  by    SEXE
        SECTION

      UNIQUE sums of squares
      All effects entered simultaneously
  
```

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	6.327	3	2.109	3.142	.026
<b>SEXE</b>	3.812	1	3.812	<b>5.679</b>	<b>.018</b>
<b>SECTION</b>	2.276	2	1.138	<b>1.696</b>	<b>.186</b>
2-Way Interactions	2.845	2	1.423	2.119	.122
<b>SEXE SECTION</b>	2.845	2	1.423	<b>2.119</b>	<b>.122</b>
Explained	10.227	5	2.045	3.047	.011
Residual	163.777	244	.671		
Total	174.004	249	.699		

250 cases were processed.  
0 cases (.0 pct) were missing.

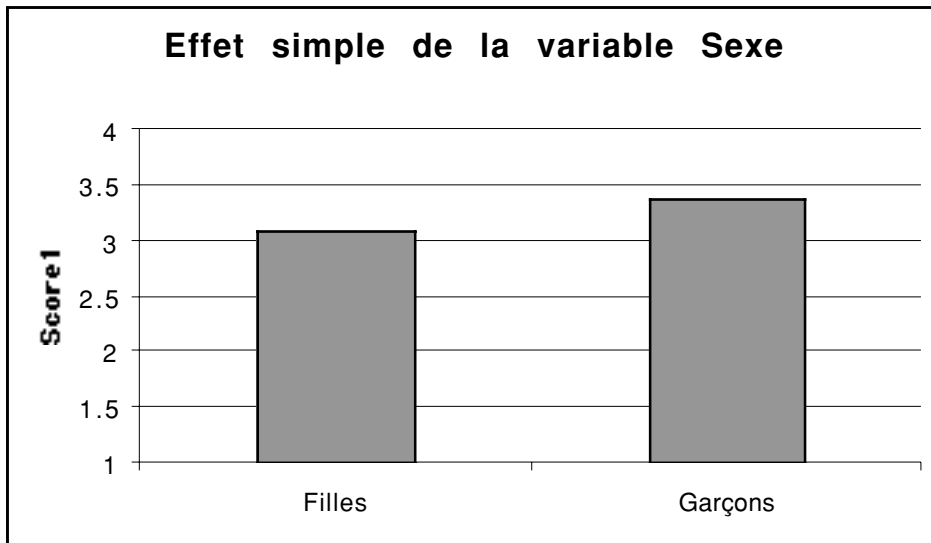


Dans la feuille des résultats, on trouve tout d'abord les effets simples ("Main Effects") pour le Sexe et la Section, puis l'effet d'interaction (s'il y a plus de deux facteurs, on trouvera plusieurs effets d'interaction deux à deux).

Les trois effets qui peuvent nous intéresser ici ont été mis en gras dans le tableau ci-dessus.

Les probabilités se trouvent dans la colonne de droite. On constate que l'effet de la variable Sexe est significative ( $p = 1.8\%$ ). Les deux autres effets (pour la Section :  $18.6\%$  ; et pour l'interaction Sexe \* Section :  $12.2\%$ ) se trouvent en dessous du seuil de  $5\%$ .

Au haut de la page suivante, un type de représentation graphique est proposé en exemple.



### Résultats pour la variable Score4

```

*** ANALYSIS OF VARIANCE ***

      SCORE4
  by    SEXE
        SECTION

      UNIQUE sums of squares
      All effects entered simultaneously

```

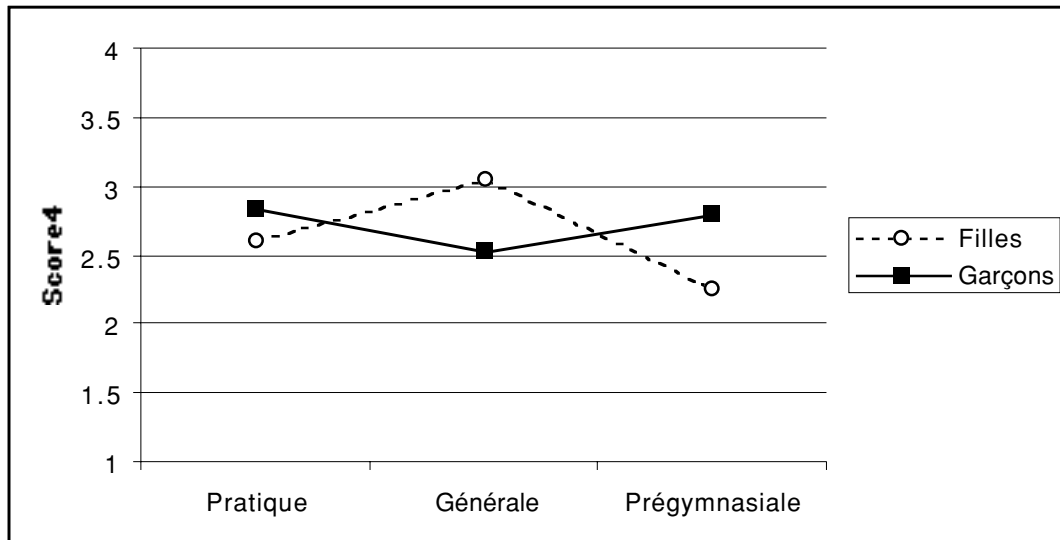
Source of Variation	Sum of Squares	DF	Mean Square	F	Sig of F
Main Effects	3.140	3	1.047	.908	.438
SEXE	.467	1	.467	<b>.406</b>	<b>.525</b>
SECTION	2.677	2	1.338	<b>1.161</b>	<b>.315</b>
2-Way Interactions	11.827	2	5.913	5.131	.007
SEXE SECTION	11.827	2	5.913	<b>5.131</b>	<b>.007</b>
Explained	14.455	5	2.891	2.509	.031
Residual	281.209	244	1.152		
Total	295.664	249	1.187		

250 cases were processed.  
0 cases (.0 pct) were missing.

Les effets qui nous intéressent ont à nouveau été mis en gras dans le tableau des résultats ci-dessus.

Il n'y a donc pas d'effets simples significatifs, par contre, la probabilité de l'effet d'interaction est inférieure à 5% ( $p = 0.7\%$ ).

On peut alors représenter graphiquement les résultats de cette manière :



Les moyennes sont calculées en utilisant les statistiques descriptives (voir pp. 9-10), séparément pour les six groupes (utilisation de la fonction "Split File").

---

## Corrélation

Les divers coefficients de corrélation permettent de tester si les variables varient ensemble.

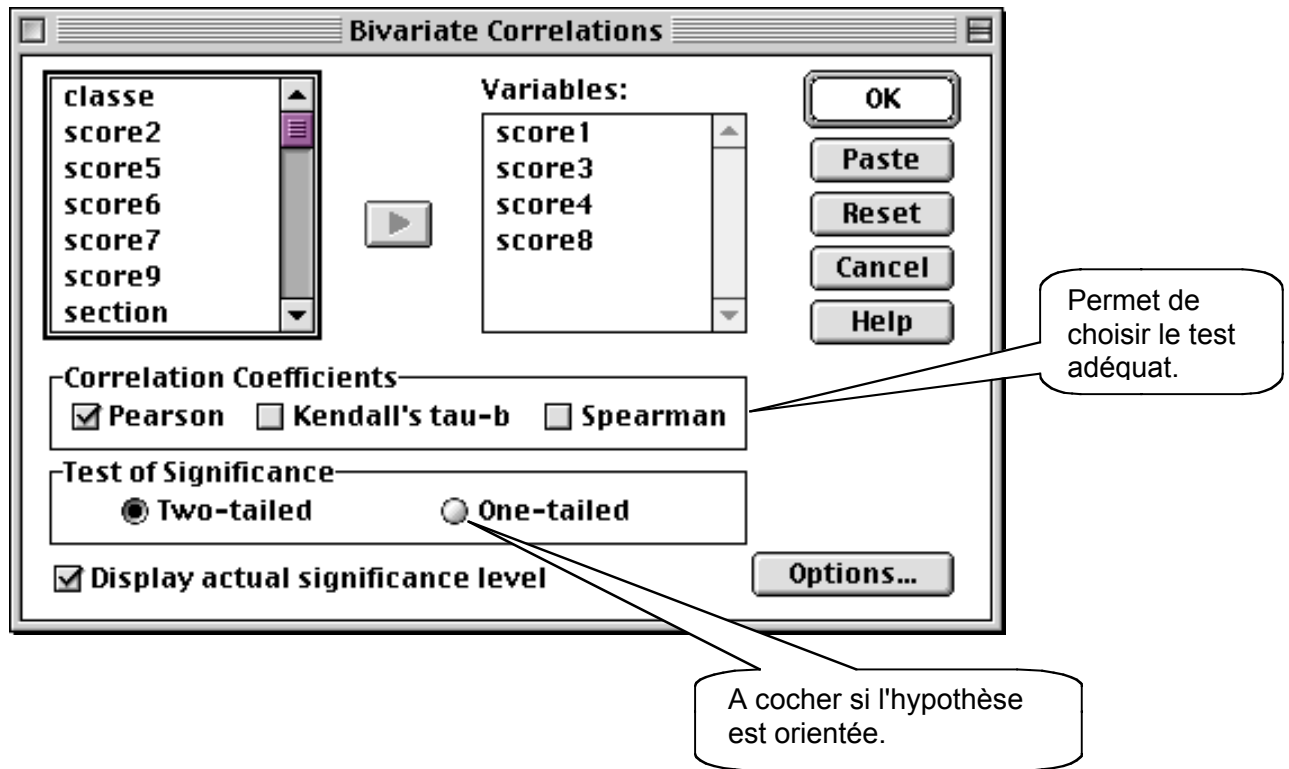
Pour les variables quantitatives, le coefficient est le  $r$  de Bravais-Pearson.

Pour les échelles ordinales, on a soit la corrélation de rang avec le  $\rho$  (rho) de Spearman, soit le  $\tau$  (tau) de Kendall.

Pour les échelles nominales, on en reste au Chi carré, avec la possibilité de calculer le coefficient de contingence ou le coefficient de Cramer.

Exemple de corrélations (deux à deux) entre les variables Score1, Score3, Score4 et Score8 (page suivante).

Chemin : Statistics → Correlate → Bivariate...



Dans notre exemple, nous ne formulons pas d'hypothèse orientée (voir explications de l'hypothèse orientée pp. 14-15), c'est-à-dire que nous postulons d'éventuelles corrélations entre les variables, sans indiquer le sens de ces liens (nous laissons donc le test à deux queues ("Two-tailed").

## Résultats

-- Correlation Coefficients --				
	SCORE1	SCORE3	SCORE4	SCORE8
SCORE1	1.0000 ( 250) P= .	.0894 ( 250) P= .159	-.1288 ( 250) P= .042	.0841 ( 250) P= .185
SCORE3	.0894 ( 250) P= .159	1.0000 ( 250) P= .	-.2354 ( 250) P= .000	.1689 ( 250) P= .007
SCORE4	-.1288 ( 250) P= .042	-.2354 ( 250) P= .000	1.0000 ( 250) P= .	-.0797 ( 250) P= .209
SCORE8	.0841 ( 250) P= .185	.1689 ( 250) P= .007	-.0797 ( 250) P= .209	1.0000 ( 250) P= .

(Coefficient / (Cases) / 2-tailed Significance)  
 " . " is printed if a coefficient cannot be computed

Coefficient de corrélation  $r$ .

Nombre de sujets.

Probabilité associée au  $r$ .

SPSS calcule toutes les corrélations possibles deux à deux et les présente sous forme de tableau croisé. On ne s'intéresse donc qu'aux résultats en dessus (ou en dessous) de la diagonale puisque ceux-ci sont identiques.

On observe donc dans nos résultats (tableau page précédente) trois corrélations significatives (mises en gras) :

Corrélation négative entre Score1 et Score4 (ce qui veut dire en d'autres termes : en général, plus le Score1 est élevé, plus le Score4 est petit).

Corrélation négative entre Score3 et Score4.

Corrélation positive entre Score3 et Score8 (ce qui veut dire en d'autres termes : en général, plus le Score3 est élevé, plus le Score8 est élevé).

On constate que si le nombre de sujets est grand (ici, 250), les corrélations seront vite significatives (dès 0.12 dans notre exemple) alors que le coefficient varie de 0 à 1 pour les corrélations positives et de 0 à -1 pour les corrélations négatives.

Si l'hypothèse est orientée (on postule à l'avance le sens de la corrélation) les valeurs des probabilités sont divisées par deux, et par conséquent, les valeurs des corrélations sont significatives à des valeurs encore inférieures à 0.12.

---

## **Régression multiple**

Les régressions multiples permettent de tenir compte de plusieurs facteurs (ou variables indépendantes) dans les relations avec la variable dépendante.

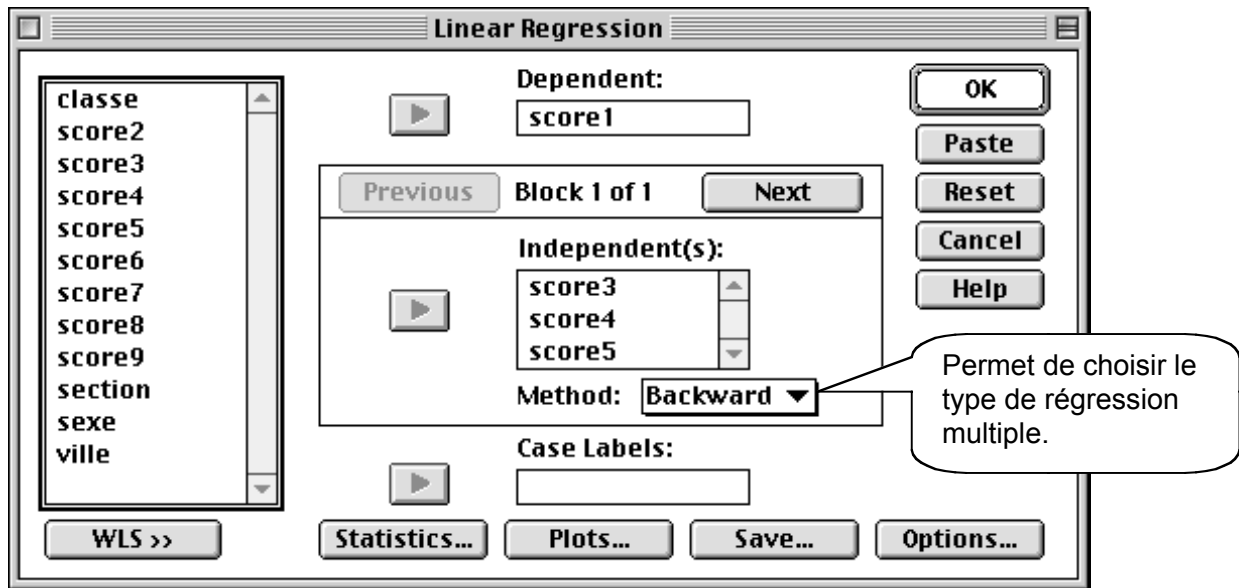
Toutes les variables doivent être quantitatives.

Il existe plusieurs modèles, nous présentons ici le modèle "Backward" qui consiste à prendre en compte l'ensemble des variables indépendantes dans la première série de calculs, puis d'enlever au fur et à mesure les variables dont la contribution est la plus faible, jusqu'à l'élaboration d'un modèle adéquat.

Les explications et l'interprétation des résultats suivront l'exemple.

Dans l'exemple qui suit, nous essayons de définir parmi les facteurs Score3, Score4 et Score5, lesquels sont en lien avec le Score1 (variable dépendante), et la force de ce lien.

Chemin : Statistics → Regression → Linear...



## Résultats

```

    * * * *  M U L T I P L E  R E G R E S S I O N  * * * *
Listwise Deletion of Missing Data
Equation Number 1   Dependent Variable..  SCORE1
Block Number 1.   Method:  Enter

Variable(s) Entered on Step Number
 1..  SCORE5
 2..  SCORE3
 3..  SCORE4

Multiple R           .20076
R Square            .04030
Adjusted R Square    .02860
Standard Error       .82391

Analysis of Variance
                DF      Sum of Squares      Mean Square
Regression        3          7.01311          2.33770
Residual         246        166.99089          .67882

F =          3.44375      Signif F =  .0174

----- Variables in the Equation -----
Variable          B          SE B          Beta          T          Sig.
SCORE3            .060095    .063031    .061274         .953    .3413
SCORE4           -.091747    .049337   -.119595        -1.860    .0641
SCORE5            .147294    .065023    .141589         2.265    .0244
(Constant)       3.086916    .238433
End Block Number  1   All requested variables entered.

```

1<sup>ère</sup> étape

Coefficient de régression multiple.

Coefficient de détermination ajusté.

Coefficient standardisé pour chaque variable.

Les trois variables sont prises en compte.

```

      * * * * M U L T I P L E   R E G R E S S I O N   * * * *
Equation Number 1   Dependent Variable..   SCORE1
Block Number  2.  Method:  Backward      Criterion  POUT  .1000
      SCORE3   SCORE4   SCORE5

Variable(s) Removed on Step Number
4..   SCORE3

Multiple R           .19172
R Square            .03676
Adjusted R Square   .02896
Standard Error      .82376

Analysis of Variance
      DF          Sum of Squares      Mean Square
Regression          2             6.39606          3.19803
Residual           247            167.60794          .67857

F =          4.71286      Signif F =   .0098

----- Variables in the Equation -----
Variable          B          SE B          Beta          T      Sig T
SCORE4            -.102827    .047940    -.134037    -2.145   .0329
SCORE5             .147832    .065009    .142106     2.274   .0238
(Constant)        3.237118    .178940
----- Variables not in the Equation -----
Variable      Beta In  Partial  Min Toler          T      Sig T
SCORE3        .061274  .060676  .943238    .953   .3413

End Block Number  2   POUT =   .100 Limits reached.

```

2<sup>ème</sup> étape

La variable avec le plus petit coefficient est enlevée du modèle.

Probabilité associée au F de Fisher.

Contributions.

Au terme de ce calcul, deux variables restent dans l'équation (Score4 et Score5), alors que Score3 a été enlevé.

On peut alors voir les contributions de chaque variable restante à l'explication de la variable dépendante (Score1).

De manière mathématique :

$$\text{Score1} = -0.103 * \text{Score4} + 0.147 * \text{Score5} + 3.24 \text{ (constante)}$$

Cette formule correspond à la formule de la droite de régression liée au coefficient de Bravais-Pearson pour les corrélations.

Les coefficients de régressions standardisés permettent de pouvoir comparer les contributions de chaque variable indépendamment de leur étendue.

## Conclusion

De nombreuses autres possibilités existent parmi les tests statistiques proposés dans *SPSS*. N'hésitez pas à utiliser le tutorial, les différentes formes d'aide à disposition, ...

Nous vous recommandons aussi d'améliorer la pertinence de vos tests en allant chercher dans les options proposées pour les différents tests, les modes de calcul ou de présentation les plus adéquats pour votre recherche.

Nous vous conseillons aussi plus spécialement les deux ouvrages suivants qui font largement référence au logiciel *SPSS*, et auxquels nous nous sommes parfois inspirés dans ce petit manuel :

Howell, D.C. (1998). *Méthodes statistiques en sciences humaines*. Bruxelles : De Boeck.

Bavaud, F., Capel, R., Crettaz de Roten, F., & Müller, J.-P. (1996). *Guide de l'analyse statistique de données avec SPSS 6*. Genève : Slatkine.

Bien entendu, les manuels en anglais (ou leurs traductions) fournies par le constructeur sont une source encore plus exhaustive, bien que parfois moins accessible à la lecture.