

## 1. Type de caractères

	Qualitatif			Quantitatif	
	nominal	dichotomique	ordinal	discret	continu
Nbre d'enfants dans un foyer				●	
Age					●
Département	●				
Note sous forme A, B, C, D, E, F			●		
Date de naissance	mauvais caractère				
Classe d'âge			●		
Note sur 20				●	
Abstentionniste		●			
Sexe	●				

On est ou n'est pas abstentionniste. Il s'agit donc bien de la présence ou l'absence d'une caractéristique. Ce n'est pas le cas du caractère Sexe. Attention donc à ne pas confondre un caractère dichotomique à un caractère nominal qui n'a que deux modalités.

## 2. Des petites questions de réflexions

- Normalement, on ne fait des calculs de moyennes que sur les caractères quantitatifs. Si on code un caractère dichotomique avec les codes 0 et 1. 0 pour l'absence de la caractéristique et 1 pour la présence de la caractéristique. Expliquer pourquoi dans ce cas précis, la moyenne des codes sur une série donnée a vraiment un sens.  
Réponse : La moyenne des codes est égale à la proportion d'individus présentant la caractéristique.
- Comment peut-on transformer un caractère quantitatif en caractère ordinal ? On différenciera les caractères discrets des caractères continus. Quelle peut être la conséquence néfaste de cette transformation.  
Réponse : Pour un caractère quantitatif discret, il n'y a rien à changer. Chaque valeur de la série devient une modalité du caractère nominal. L'ensemble R étant ordonné, le caractère est donc ordinal. Pour un caractère quantitatif continu, il faut le transformer en définissant un nombre fini d'intervalles contigus et disjoints. Inconvénient : on perd de l'information.
- Pourquoi il est à priori incorrect de transformer un caractère ordinal en un caractère quantitatif discret ?  
Réponse : avec un caractère quantitatif, certains indicateurs sont obtenus en faisant des différences entre les valeurs. Si on transforme les modalités d'un caractère ordinal par des nombres on fait l'hypothèse (forte) que les écarts entre les modalités successives sont égaux aux écarts des nombres correspondants.

4. Rappeler pourquoi quand on a établi une échelle de richesse d'information entre les différents types de caractères.

Réponse : un caractère peut toujours être transformé en un caractère d'un type moins riche. En d'autres termes, tous les traitements prévus pour un type peuvent être utilisés pour tout caractère d'un type plus riche.

### 3. Echantillonnage par strate

1. On étudie un caractère  $X$  sur une population de grande taille. Des études précédentes ont montré que ce caractère est très lié à deux autres caractères nominaux  $Y$  et  $Z$ . On désire utiliser l'échantillonnage par strate en définissant les strates à l'aide des deux caractères  $Y$  et  $Z$ .

- a. Quelles sont les informations dont on doit disposer pour réaliser cette stratification

Réponse : pour chaque couple de modalités  $(y_i, z_j)$ , on doit disposer de la proportion des individus qui présentent simultanément la modalité  $y_i$  pour  $Y$  et la modalité  $z_j$  pour  $Z$ .

- b. Quels peuvent être les inconvénients (voire les impossibilités) de cette stratification avec deux caractères.

Réponse : On multiplie le nombre de strates. Il faut donc que l'échantillon soit suffisamment grand pour que l'effectif par strates soit suffisant.

2. Ecrire l'algorithme qui permet de construire un échantillon par strate. On suppose qu'on que les individus de la population sont connus et numérotés de 1 à  $N$ . Pour chaque individu, on connaît sa strate d'appartenance.

#### 4. Un petit cas

Dans un collège, on a deux sixièmes : 6<sup>ème</sup> A (21 élèves) et 6<sup>ème</sup> B (21 élèves). Ces collégiens passent un test de maths. Les notes obtenues sont les suivantes :

	6 <sup>ème</sup> A	6 <sup>ème</sup> B
1	10	9
2	11	9
3	9	12
4	9	12
5	13	17
6	7	12
7	8	12
8	13	18
9	9	8
10	15	13
11	10	8
12	11	12
13	8	18
14	12	8
15	11	8
16	14	17
17	14	14
18	9	14
19	12	8
20	12	10
21	10	12

1. Faire sur un même graphique les diagrammes en bâton de chacune des séries de notes.
2. Faire pour chaque classe la boîte à moustache de la série de notes.
3. Interpréter les résultats par classe et en comparant les classes.