

# Chapitre III

## La conception du questionnaire

---

Cette étape est étape clef, elle dépend étroitement des étapes différentes (surtout l'objectif de l'étude)

Questionnaire : ensemble de questions construit dans le but de générer l'information nécessaire à l'accomplissement des objectifs de l'étude.

L'idée d'un questionnaire est un outil de standardisation de l'information collectée. On pose la même question à des individus on peut donc comparer sans risque les différentes réponses.

### I) TRAVAIL DE FOND DU QUESTIONNAIRE

#### 1) Le contenu des questions

Au préalable établir une liste des informations à rechercher avant de poser les questions. Cette liste est basée sur les objectifs de l'étude et sur les hypothèses de travail qui ont été fixées. Le questionnaire sera donc créé à parti de cette liste. Grace à cette liste il n'y a aucun oubli et les questions ne sont pas inutiles.

On s'intéresse à la nature des informations que l'on va rechercher. (Questions en termes de comportements → fréquence d'achat ou des questions de données sociogéographique, ou données psychologique → style de vie, personnalité...)

Avant de placer une question dans le questionnaire, elle doit être soumise à 4 interrogations :

- Cette question est-elle strictement nécessaire ? Est-ce-que un des objectifs de l'étude sans cette question ?
- Pour l'information recherchée, une seule question suffira-t-elle ?
- La personne interrogée pourra-t-elle fournir l'information ? (question soit un peu trop pointue, soit un peu trop général)

Il y'a 3 cas où la personne ne fournit pas l'information :

- ☆ Incompréhension (domaine trop technique)
- ☆ Incompréhension abstraction trop compliquée

- ☆ Ignorance, elle traduit une mauvaise adaptation du questionnaire par rapport à la personne interrogée, ou alors une erreur dans la population par rapport à l'information
- ☆ Oublis, la personne n'est pas capable de la restituer (combien de yaourt avez-vous mangé les 6 derniers mois ?)
- La personne voudra-t-elle fournir l'information ? il existe des informations sensibles (vie privée). Essayer de mettre les questions sensibles vers la fin. On évite de demander les choses trop brutes (trancher de revenus au lieu du revenu brut)

## 2) Les différents types de questions

### a) Les questions fermées

Les personnes interrogées doivent choisir une ou plusieurs réponse(s) parmi des réponses formulées à l'avance par le rédacteur du questionnaire.

#### **Les avantages de ces questions fermées :**

- ☆ Elles facilitent la compréhension de la question
- ☆ Elles facilitent l'expression de la réponse
- ☆ Elles facilitent l'analyse

#### **Les inconvénients :**

- ☆ Elles limitent les possibilités du répondant
- ☆ Elles sont difficiles à rédiger

#### **Les différents types de questions fermées :**

- ☆ Questions fermées dichotomiques : la personne a deux solutions (oui ou non) → son avantage est : simplicité du questionnement et du traitement. Son inconvénient il n'y pas d'autres alternatives soit oui soit non
- ☆ Questions à choix multiples (QCM) : la personne a la liberté de choisir une ou plusieurs réponses.
- ☆ Question multichotomique à réponse unique : la personne ne peut cocher qu'une réponse.
- ☆ Questions multichotomique à réponse multiples : proposition d'une liste soit la personne peut choisir autant qu'elle veut ou seulement un certains nombre de réponse.
- ☆ Question avec classement hiérarchique : on demande à la personne d'ordonner la liste des propositions qu'on lui demande (soit ordonner tous les éléments soit ordonner les 3 premiers)
- ☆ Questions avec échelle d'attitude : l'intérêt est de quantifier les informations et d'avoir des nuances. Il existe différents types d'échelles.

### ***Echelle Likert :***

Elle a pour but de mesurer l'attitude d'un individu en additionnant les notes que l'individu attribue à une liste de proposition. Cette échelle contient 5 niveaux qui permettent de nuancer le degré d'accord → on traduit toutes les dimensions d'un problème en affirmation et on demande aux personnes de se situer

5 niveaux

- Pas du tout d'accord
- Pas d'accord
- Neutre
- D'accord
- Tout à fait d'accord

Exemple :

Pour moi protéger l'environnement cela consiste à :

- Utiliser des lessives sans phosphates
- Trier les déchets ménagers
- Acheter des produits bio.

Certains auteurs pensent que l'échelle de Likert doit être en nombre de pair, car cela les oblige à sélectionner, il n'y a pas de neutre, ils doivent se positionner. Ainsi les répondants n'ont pas la tendance à se positionner au milieu (au neutre). Cette échelle permet de mesurer le degré de satisfaction.

### ***Echelle différentielle sémantique ou d'Osgood :***

Très utilisée pour l'image. C'est une échelle dont les 2 bornes extrêmes sont des adjectifs de sens opposé. Le nombre d'échelons proposés est généralement de 5 ou 7 :

Ex : « qu'est ce que vous pensez de notre entreprise :

- Moderne/ dépassée
- Active/ inactive
- Malhonnête/ honnête.

Plusieurs présentations possibles :

- Cocher une case
- Soit proposer des adjectifs

- Proposer un axe (segment) et demander à la personne de mettre la croix là où elle veut.

### ***Les échelles d'intention :***

L'objectif est de mesurer une intention déclarée du comportement.

Exemple : « Si la Fnac s'ouvrait à Nouméa ? »

- J'achèterai mes CD à la Fnac/ je n'achèterai pas à la Fnac

Il existe d'autres échelles :

- Echelles figuratives (smiley face)

### b) Les questions ouvertes

Le sondé répond librement à la question, il n'est aucunement influencé par le choix de réponse. Cependant, il y'a beaucoup d'inconvénients :

- Beaucoup de non réponse
- Réponse inexploitable
- Administration est assez difficile (l'auteur doit tout noter)
- Le tri et l'analyse de ces questions sont délicats et nécessitent un logiciel spécialisé

### ***Il existe différents types :***

- Questions ouvertes numériques → facilement traitable → elles permettent de connaître des ordres de grandeurs, des niveaux de dépenses, de quantités, de fréquences... il faut éviter des thèmes trop confidentiels ou demandant trop d'effort.  
« Quelle note sur 10 donneriez-vous à l'accueil téléphonique ? »  
« Combien d'heures par semaines utilisez-vous internet par semaine ? »  
Les avantages → Plus riche, plus précis qu'une échelle ou que les classes que l'on peut proposer. → Maximum de choix et plus le taux de réponse est élevé.  
L'inconvénient → Complexité du traitement et parfois les personnes peuvent ne pas savoir quelle réponse donnée ou de ne pas vouloir.
- Questions ouvertes à texte :

On est souvent tenté d'utiliser de type de question → connaître la motivation ou le frein d'utiliser un produit. Ces questions peuvent être transformées en une série de questions fermées (plus faciles à traiter)

## II) LA FORME DU QUESTIONNAIRE

### 1) La rédaction du questionnaire

Le vocabulaire et le style utilisé pour la rédaction du questionnaire doit être employé par les répondants et non les rédacteurs.

#### **Quelques conseils :**

- Tout vocabulaire technique (sauf cas particulier) est à bannir (ou on le définit)
- Tout vocabulaire ambigu est à bannir (confusion, compréhension subjective)
- La perception du temps (tout les mots se terminant en « ant » doivent être évités : fréquemment, souvent, de temps en temps)
- Eviter les questions trop longues et trop complexes

☆ Vocabulaire simple et bien adapté à la population

### 2) La structure du questionnaire

Le questionnaire s'organise en 7 étapes (technique de l'entonnoir)

- ☆ Présentation de l'enquêteur : on se présente (on saluer l'interlocuteur) et on présente en quelques phrases le thème du questionnaire. → introduction rédigée
- ☆ Question d'introduction : c'est le seul endroit dans le questionnaire où l'information peut ne pas être utile, c'est pour placer le répondant dans de bonnes conditions (questions générales assez faciles où les réponses sont souvent positives)
- ☆ Des questions qualifiantes : ces questions n'ont lieu d'être que si on a besoin de savoir si l'individu que l'on interroge possède ou non l'information. Elles servent d'aiguillage : « Possédez-vous un compte bancaire » = oui/non. Si non on peut faire un renvoi sur le questionnaire.
- ☆ Questions de mise en route : Elles servent à centrer précisément le thème de l'étude en commençant par les questions simple (questions de comportement → combien de yaourt consommez-vous par semaine ?)

- ☆ Questions spécifiques (dans le cœur du questionnaire) : ces questions amènent les informations les plus importantes. A ce stade le répondant est imprégné du thème. On peut demander à ce moment là , un effort de réflexion et/ou demander des informations un plus personnelles (motivations, opinion, jugement personnel)
- ☆ Les questions d'identification : tout ce qui sert à décrire le profil de la personne interrogée (profession, sexe, lieu de résidence, âge...)
- ☆ On remercie la personne, donnez un site ou un endroit où les résultats sont affichés.
- ☆ Question piège : vérifier si la personne est honnête. (2 questions pareil mais changé le vocabulaire)

### III) L'ADMINISTRATION

#### 1) Le pré-test

Il consiste à administrer le questionnaire à un petit nombre d'individus appartenant à la population ciblée par le questionnaire.

Son but est de détecter les erreurs commises lors de la rédaction du questionnaire → méthode est la même pour l'échantillonnage. Les principales erreurs sont : un mauvais vocabulaire ou de style → les personnes ne comprennent pas tout, demandent des informations en plus.

#### ***Un mauvais choix des modalités de réponses :***

- ☆ Si 80% de réponses sont pour une seule modalité (surtout la modalité « autre ») il faut revoir les modalités, préciser les modalités ou ajouter les modalités.
- ☆ Erreurs dans les questions filtres
- ☆ Manque de clarté dans les instructions (questionnaire auto-administré) il faut que la personne comprenne.

### IV) LES TRAITEMENTS DES DONNÉES

Les analyses peuvent être plus au moins complexe.

#### 1) Le tri à plat

Le tri à plat consiste à réorganiser l'ensemble des valeurs prises par une seule variable. Les variables peuvent être :

- variables qualitatives, elles ne sont pas mesurables, les modalités sont plutôt des qualités de l'objet étudié. Ex : la couleur des yeux. Elles sont nominales ou ordinales (on peut demander un ordre de classement)
- Variables quantitatives, elles sont mesurables par une unité de mesure. Elles sont métriques (on peut donner un ordre et on connaît la distance entre deux variables). Il y a des variables discrètes → variables ne prenant que certaines valeurs (nombre d'enfant)  
Les variables continues

a) Les variables qualitatives

Variable : sexe	Modalité ( $X_i$ )	Effectif ( $N_i$ )	Fréquence ( $f_i$ )
Hommes	1	9	0,45
Femmes	2	11	0,55
	Total	N=20	1

**NE PAS FAIRE DES MOYENNES CAR IL S'AGIT DE VARIABLES QUALITATIVES !!**

b) Variables quantitatives discrètes

Pour ces variables il n'y a pas de difficultés majeures, on peut effectuer un calcul de la fréquence relative, fréquence cumulée, fréquence croissante ou décroissante.

Variables :	Modalités : $X_i$	Effectifs : $n_i$	Fréquence : $f_i$
utilisation d'une lessive sans phosphates			
Pas du tout d'accord	1	1	0,05
Plutôt pas d'accord	2	3	0,15
Sans opinion	3	3	0,15

Plutôt d'accord	4	5	0,25
Tout à fait d'accord	5	8	0,40
	total	N = 20	1

c) *Les variables quantitatives continues*

Il faut regrouper les variables en classe. Ce regroupement peut se faire en amont ou après. Puis créer des centres de classe.

2) *Les tris croisés*

Quand on croise le valeurs de 2 variables, ce qui revient à construire un tableau a double entrée (tableau de contingence).

	18_24 ans	25-34	35-44	45-60	>60	TOTAL
Hommes	0 (0%) (0%)	3(33%) (43%)	1(11%) (25%)	5(56%) (83%)	...	...
Femmes	1 (9%) (100%)	4(36%) (57%)	3(27%) (75%)	1(9%) (17%)	...	...
total	1(5%)	7(35%)	4(20%)	6(30%)	...	...

3) *Les calculs statistiques de base*

Ces calculs permettent de repérer la tendance centrale (la moyenne, la médiane → 50% au dessus et 50% en dessous, et le mode), la dispersion d'une variable.

a) *La moyenne*

C'est la mesure la plus centrale courante, on ne la calcule que sur les variables quantitatives discrètes et continues.

- N observation regroupées selon k modalités : (variable discrète)

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k x_i n_i$$

I = 1, ... k

K= nombre de modalités

$X_i$  = ième modalité

$N_i$ = effectif de la ième modalité

- N observation regroupées selon c classes : (variable continue)

$$\bar{x} = \frac{1}{n} \sum_{i=1}^c x_i n_i$$

C = nombre de classes

$X_i$  = centre de la ième classe

$N_i$ = effectif de la ième classe

### Les limites de la moyenne :

- Elles sont métriques
- Elles sont influencées par les valeurs extrêmes. Il ne faut pas se limiter seulement à la moyenne mais aux valeurs de distribution
- Ce n'est qu'un paramètre d'une tendance centrale, il faut la compléter par une analyse par dispersion.

### b) Variance et écart-type

La variance permet de nous donner pour chaque point entre la moyenne et le point.

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^k n_i (x_i - \bar{x})^2 = \frac{1}{n} \left( \sum_{i=1}^k n_i x_i^2 \right) - \bar{x}^2$$

Ecart-type :  $\sigma = \sqrt{\sigma^2}$

L'écart-type est l'écart de la moyenne.

Le coefficient de variation :  $CV = \frac{\sigma}{\bar{x}}$

Il permet de comparer les différentes données car il permet d'enlever les unités de mesure. Quand le coefficient est supérieur à 25%, la est dispersée.

4) *Les tests d'ajustement et de contingence*

a) *Les tests d'ajustement chi-deuse ( $X^2$ )*

Exemple :

PCS	Cod e	Effectifs dans la population	Fréquenc e dans la population	Effectif observé dans l'échantillo n	Effectif théorique dans l'échantillo n	$X^2$
Agriculteurs exploitants	1	4316	0,0143	3	5,59	1,20001789
Artisans-commerçants	2	15044	0,0499	33	19,51	9,32752947
Professions supérieures	3	13120	0,0435	65	17,02	135,257368
Professions intermédiaires	4	33636	0,1116	23	43,63	9,75468485
Employés	5	43436	0,1441	90	56,34	20,1099681
Ouvriers	6	48810	0,1619	24	63,32	24,4166519
Retraités	7	54276	0,1801	33	70,41	19,876553
étudiants	9	57814	0,1918	95	75	5,33333333
Sans emploi+mères au foyer	8+10	30968	0,1027	25	40,17	5,72887478
Total		301 420	1,0000	391	391	231,004981

$391 \cdot 0,0143 = 5,59$  : j'aurais du interrogé 5,59 agriculteurs exploitants (donc environ 6)

$X^2 = (\text{Calculé effectif observé} - \text{effectif théorique})^2 / \text{théorique} = (3 - 5,59)^2 / 5,59 = 1,2$

$\alpha = 5\%$                        $k=9$  (modalité)      degré de liberté : nombre de modalité en colonne  
 $-1 : v = 9-1=8$

$$= \frac{(3-5,6)^2}{5,6} + \frac{(33-19,51)^2}{19,51} + \dots + \frac{(25-40,17)^2}{40,17} = 231$$

$X^2_{\text{calculé}} = 231$                        $X^2_{\text{dans la table}} = 15,5$

$X^2_{\text{calculé}} > X^2_{\text{table}} \rightarrow$  on rejette  $H_0$        $X^2_{\text{calculé}} = X^2_{\text{table}} \rightarrow$  on accepte  $H_0$  si la différence entre les deux  $X^2$  est assez faible on peut accepter  $H_0$

$H_0 \rightarrow$  mon échantillon est représentatif de la population

Si on rejette  $H_0 \rightarrow$  mon échantillon n'est pas représentatif de la population (si j'interroge les gens à une certaines heures)

b) les tests de contingence

Destination choisie										
	Baléares	Roches	Egypte	Total	Efficatif théorique (ET) Baléares	ET Roche	ET Egypte	$X^2$ Baléares	$X^2$ Roches	$X^2$ Egypte
Primaire	300	50	100	450	270	112,5	67,5	3,3333 3333	34,722 2222	15,648 1481
Secondaire	250	80	20	350	210	87,5	52,5	7,6190 4762	0,6428 5714	20,119 0476
Supérieur	50	120	30	200	120	50	30	40,833 3333	98	0
Total	600	250	150	1000	600	250	150	51,785 7143	133,36 5079	35,767 1958

Somme  $X^2 = 51,7857143 + 133,365079 + 35,7671958 = 220,917989$

$1000 * 0,6 * 0,45 = 270$

Degré de liberté : (nombre de modalité en colonne -1) \* (nombre de modalité en ligne -1) =  $(3-1) * (3-1) = 4$

$X^2_{\text{table}} = 9,488$

$$X^2_{\text{calculé}} > x^2_{\text{table}} = 220,91 > 9,488$$

$H_0: X^2_{\text{calculé}} = X^2_{\text{table}} \rightarrow$  les 2 variables sont indépendantes

$H_1: X^2_{\text{calculé}} \neq X^2_{\text{table}} \rightarrow$  il y'a un lien entre les 2 variables

Ainsi je rejette  $H_0$ . J'aurais dû interroger 270 personnes voulant aller aux Baléares et étant au niveau primaire.

On peut aussi utiliser covariance, corrélation et analyse des droites de régression.