# Échantillonnage

## Introduction : Pourquoi sonder ?

**Exercices corrigés**

---

Jaouad Madkour
Faculté d'Économie et de Gestion de Tanger
Échantillonnage

## Table des matières

Partie 1 — Énoncés . . . . . 1

Partie 2 — Corrections . . . . . 3

# 1 Exercices du chapitre

### Exercice 1 — Recenser ou sonder ?

Pour chacune des situations suivantes, dire s'il faut recenser la population ou la sonder, et justifier en une phrase.

(a) Connaître le nombre exact d'habitants de chaque commune, pour répartir des dotations budgétaires.

(b) Estimer la durée de vie moyenne d'un lot de 10 000 ampoules sortant d'une usine.

(c) Mesurer chaque trimestre le taux de chômage national.

(d) Connaître l'avis des 45 salariés d'une PME sur un projet de réorganisation.

### Exercice 2 — Nommer les éléments d'une enquête

La régie de distribution de Tanger veut estimer la consommation d'eau mensuelle moyenne des ménages de la ville. Elle tire au hasard 800 compteurs dans son fichier d'abonnés et relève leur consommation.

(a) Préciser la population, l'unité, la base de sondage, l'échantillon et le paramètre.

(b) Citer deux situations dans lesquelles un compteur ne correspond pas exactement à un ménage. Quelle erreur en résulte ?

(c) Des ménages non raccordés s'alimentent à des bornes-fontaines. Comment s'appelle l'erreur qu'ils provoquent ? Dans quel sens risque-t-elle de fausser la moyenne ?

### Exercice 3 — Paramètre, estimateur, estimation

Six ménages ont des dépenses mensuelles de 5, 7, 8, 10, 12 et 18 milliers de dh. Un enquêteur tire trois ménages au hasard et calcule la moyenne de leurs dépenses.

(a) Calculer le paramètre $\overline{Y}$.

(b) L'enquêteur tire les ménages dépensant 7, 12 et 18. Quelle estimation obtient-il ? Un collègue tire 5, 8 et 10 : quelle est la sienne ?

(c) Combien d'échantillons de trois ménages pouvait-on tirer ?

(d) Lequel des trois objets — paramètre, estimateur, estimation — est aléatoire ? L'un des deux enquêteurs s'est-il trompé ?

# 2 Applications en sciences économiques

### Exercice 4 — Le budget d'une enquête communale

Une commune compte 40 000 ménages. Une enquête coûte 200 000 dh de frais fixes (conception, formation, saisie) et 150 dh par ménage interrogé.

(a) Calculer le coût d'une enquête exhaustive et celui d'un sondage de 1 200 ménages.

(b) Pour une proportion voisine de 0,5, la marge d'erreur à 95 % d'un sondage aléatoire simple vaut $e = 1{,}96\,\sqrt{(1 - n/N) \times 0{,}25/n}$. La calculer pour $n = 1\,200$.

(c) La commune hésite à consacrer 16 fois plus d'argent pour supprimer cette marge d'erreur. Quel argument lui opposer, en pensant aux erreurs de non-échantillonnage ?

### Exercice 5 — Un sondage par annuaire

Dans une ville, 30 % des ménages ont une ligne téléphonique fixe. Parmi eux, 90 % ont accès à internet, contre 55 % parmi les autres. Un institut interroge 20 000 ménages tirés dans l'annuaire des lignes fixes

pour estimer la part des ménages connectés.

(a) Quelle est la vraie proportion de ménages connectés dans la ville ?

(b) Autour de quelle valeur se situera l'estimation de l'institut ? Sa marge d'erreur d'échantillonnage vaut $\pm 0{,}4$ point : est-ce rassurant ?

(c) Nommer l'erreur commise et proposer une base de sondage plus adaptée.

# 1 Exercices du chapitre

### Correction — Exercice 1

(**a**) Recenser : on veut un chiffre exact pour **chaque** commune, y compris les plus petites, où un échantillon serait trop réduit.
(**b**) Sonder : mesurer la durée de vie détruit l'ampoule ; tester tout le lot ne laisserait rien à vendre.
(**c**) Sonder : un recensement chaque trimestre serait hors de prix, et le résultat arriverait trop tard.
(**d**) Recenser : avec 45 personnes, interroger tout le monde ne coûte presque rien de plus qu'un échantillon, et supprime l'erreur d'échantillonnage.

### Correction — Exercice 2

(**a**) Population : les ménages résidant à Tanger. Unité tirée : le compteur (unité d'observation : le ménage). Base : le fichier des abonnés. Échantillon : les 800 compteurs tirés. Paramètre : la consommation mensuelle moyenne d'eau des ménages de Tanger.
(**b**) Un immeuble dont plusieurs ménages partagent un compteur collectif ; un ménage qui a deux compteurs (logement et local). Le compteur n'est pas l'unité visée : c'est une erreur de **couverture** (doubles comptes ou ménages regroupés), qui fausse la consommation « par ménage ».
(**c**) Une erreur de **couverture** : ces ménages ne peuvent pas être tirés. Ils consomment probablement moins d'eau du réseau que les autres : la moyenne des abonnés surestime la consommation moyenne de l'ensemble des ménages.

### Correction — Exercice 3

(**a**) $\overline{Y} = (5+7+8+10+12+18)/6 = 60/6 = 10$ milliers de dh.
(**b**) $\overline{y} = (7+12+18)/3 = 12{,}33$ pour le premier ; $(5+8+10)/3 = 7{,}67$ pour le second.
(**c**) $\binom{6}{3} = 20$ échantillons possibles.
(**d**) Le paramètre (10) est fixe. L'estimateur — la règle « moyenne des trois dépenses » — est une variable aléatoire, puisque sa valeur dépend de l'échantillon tiré ; chaque estimation en est une réalisation. Aucun des deux ne s'est trompé : ils ont appliqué la même règle à deux échantillons différents. L'un tombe au-dessus, l'autre au-dessous ; le chapitre 2 montrera qu'en moyenne sur les 20 échantillons, la règle retrouve exactement 10.

# 2 Applications en sciences économiques

### Correction — Exercice 4

(**a**) Exhaustive : $200\,000 + 150 \times 40\,000 = 6\,200\,000$ dh. Sondage : $200\,000 + 150 \times 1\,200 = 380\,000$ dh, soit 16 fois moins.
(**b**) $e = 1{,}96\sqrt{(1 - 1\,200/40\,000) \times 0{,}25/1\,200} = 1{,}96 \times \sqrt{0{,}97 \times 0{,}000208} = 0{,}0279$, soit $\pm 2{,}8$ points.
(**c**) L'enquête exhaustive supprime l'erreur d'échantillonnage, mais pas les autres : 40 000 entretiens demandent beaucoup plus d'enquêteurs, moins formés et moins contrôlés, et donc davantage d'erreurs de mesure et de non-réponse. Un sondage de 1 200 ménages bien conduit peut être plus juste au total.

### Correction — Exercice 5

(**a**) $P = 0{,}30 \times 0{,}90 + 0{,}70 \times 0{,}55 = 0{,}27 + 0{,}385 = 0{,}655$, soit 65,5 %.
(**b**) L'institut n'interroge que des ménages ayant une ligne fixe : son estimation tournera autour de 90 %, à $\pm 0{,}4$ point près. La marge est minuscule, mais elle mesure la dispersion autour de 90 %, pas autour de la vraie valeur : l'estimation est fausse de 24,5 points, très précisément.
(**c**) Une erreur de **couverture** (base de sondage biaisée), exactement celle du Literary Digest. Il faut

une base couvrant tous les ménages : la liste des ménages du recensement, mise à jour, ou un tirage de logements sur le terrain.