

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET  
DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITÉ DE CARTHAGE

---

ÉCOLE SUPÉRIEURE DE LA STATISTIQUE  
ET DE L'ANALYSE DE L'INFORMATION



---

Projet de Sondage : Enquête sur les  
dépenses en friperie

---

Élaboré par : Bellalah Yesmine  
Jaouadi Nour  
Madani Sarra

Encadré par : Mme Mallek Hela  
Année Universitaire 2016/2017

---

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Partie Descriptive</b>	<b>2</b>
2.1	Principales étapes d'une enquête . . . . .	2
2.2	Questionnaire . . . . .	2
2.3	Collecte de données . . . . .	2
2.4	Description de la base de données . . . . .	2
2.5	Étude descriptive . . . . .	3
<b>3</b>	<b>Partie Analytique</b>	<b>12</b>
3.1	Sondage à probabilités égales . . . . .	13
3.1.1	Sans remise(PESR) . . . . .	13
3.1.2	Avec remise (PEAR) . . . . .	15
3.2	Sondage à probabilités inégales . . . . .	17
3.2.1	Avec remise PIAR . . . . .	17
3.2.2	Sans remise PISR . . . . .	19
3.3	Startification . . . . .	21
3.3.1	Strates à tailles égales . . . . .	21
3.3.2	Strates à allocations proportionnelles . . . . .	23
3.3.3	Stratification à allocation optimale . . . . .	26
3.4	Sondage par grappes et à deux degrés . . . . .	29
3.4.1	Sondage par grappe . . . . .	29
3.4.2	Sondage à deux degrés . . . . .	31
<b>4</b>	<b>Synthèse</b>	<b>33</b>
<b>5</b>	<b>Comparaison et commentaires</b>	<b>34</b>
<b>6</b>	<b>Redressement</b>	<b>35</b>
6.1	Redressement par post stratification . . . . .	35
6.2	Redressement par le quotient . . . . .	36
<b>7</b>	<b>Conclusion</b>	<b>37</b>
<b>8</b>	<b>Annexe</b>	<b>38</b>

---

## Table des figures

1	Questionnaire . . . . .	3
2	Camembert de la variable sexe . . . . .	4
3	Camembert de la variable age . . . . .	4
4	Camembert de la variable budget . . . . .	5
5	Camembert de la variable catégorie socio-professionnelle . . . . .	6
6	Camembert de la variable état matrimonial . . . . .	6
7	Camembert de la variable niveau d'étude . . . . .	7
8	Camembert de la variable préférence d'achat à la friperie . . . . .	7
9	Budget moyen en fonction de la variable sexe . . . . .	8
10	Budget moyen en fonction de la variable age . . . . .	8
11	Budget moyen en fonction de la variable niveau d'étude . . . . .	9
12	Budget moyen en fonction de la variable catégorie socio professionnelle . . . . .	9
13	Budget moyen en fonction de la variable niveau d'étude . . . . .	10
14	Budget moyen en fonction de la variable préférence . . . . .	11
15	Corrélation des variables auxiliaires avec la VI . . . . .	11
16	Matrice de corrélation . . . . .	12
17	Matrice d'échantillons PESR . . . . .	13
18	Matrice des estimations PESR . . . . .	14
19	Graphe des intervalles de confiance PESR . . . . .	14
20	Matrice d'échantillons PEAR . . . . .	15
21	Matrice des estimations PEAR . . . . .	16
22	Graphe des intervalles de confiance PEAR . . . . .	16
23	Matrice d'échantillons PIAR . . . . .	17
24	Matrice des estimations PIAR . . . . .	18
25	Graphe des intervalles de confiance PIAR . . . . .	18
26	Matrice d'échantillons PISR . . . . .	19
27	Matrice des estimations PISR . . . . .	20
28	Graphe des intervalles de confiance PISR . . . . .	20
29	Matrice d'échantillons issus de la stratification à taille égale . . . . .	21
30	Matrice des estimations par stratification à taille égale . . . . .	22
31	Graphe des intervalles de confiance par la stratification à taille égale . . . . .	23
32	Matrice d'échantillons à allocation proportionnelle . . . . .	24
33	Matrice des estimations par stratification à allocation proportionnelle . . . . .	25
34	Graphe des IC0(0.95) par stratification à allocations proportionnelles . . . . .	26
35	Matrice d'échantillons à allocation optimale . . . . .	27
36	Matrice des estimations à allocation optimale . . . . .	28
37	Graphe des intervalles de confiance à allocations optimales . . . . .	29
38	Matrice d'échantillons issus de la méthode par grappe . . . . .	30
39	Matrice des estimations par la méthode des grappes . . . . .	30
40	Graphe des intervalles de confiance par la méthode des grappes . . . . .	31
41	Matrice des estimations par le sondage à deux degrés . . . . .	32

---

42	Graphe des intervalles de confiance par le plan à deux degrés . . . . .	32
43	Les différents Intervalles de confiance . . . . .	33
44	Matrice des moyennes et des IC pour chaque méthode de sondage . . . . .	34
45	Boîte à moustaches des longueurs IC0.95 par plan de sondage . . . . .	34
46	Les estimations par post stratification . . . . .	36
47	Les estimations par redressement par le quotient . . . . .	36

---

# 1 Introduction

Un sondage est une méthode consistant à recueillir des informations sur un groupe d'individus ou d'unités statistiques (Échantillon) prélevé dans un ensemble plus important (Population), et destinée à estimer certaines caractéristiques de la distribution d'un ou de plusieurs caractères dans cette population dans un but soit descriptif, soit décisionnel.

Un ensemble de théories appelé « Théorie de sondage » est formé autour de ce concept statistique. Cette dernière nous renseigne sur les fondements mathématiques et méthodologiques des enquêtes par sondage afin de fournir une base de données fiables et pertinentes sur tout ce qui concerne la taille, la composition et les caractéristiques de la population cible. C'est dans cette optique que nous avons réalisé une enquête sur les dépenses en friperie.

Notre variable d'intérêt (VI) est le budget moyen consacré pour la friperie pendant les 3 derniers mois.

L'objectif visé par cette enquête d'estimer la VI par application des différentes méthodes de sondage vues en classe afin d'aboutir à une comparaison entre ces différentes méthodes.

Dans ce rapport on commence en première partie par une étude descriptive (Définitions et statistiques descriptives relatives aux variables). En seconde partie on procède aux résultats des opérations d'échantillonnage et d'estimation (Ponctuelle et par intervalle de confiance).

Le rapport s'achève par une conclusion générale sur l'enquête ainsi que les études effectuées.

---

## **2 Partie Descriptive**

### **2.1 Principales étapes d'une enquête**

- Définir les objectifs des enquêtes en fixant les vraies valeurs qu'on cherche à estimer.
- Définir la population tout en déterminant la base de sondage adéquate.
- Tirage de l'échantillon.
- Préparation du questionnaire.
- Collecte des données et les enregistrer.
- Contrôle des données.
- Codage des données (Chiffrement).
- Saisie des données.
- Procéder aux estimations et analyse des données.
- Publication des résultats.

### **2.2 Questionnaire**

Le questionnaire est l'une des grandes méthodes pour le recueil de données. C'est une méthode de collecte des informations en vue de comprendre et d'expliquer les faits. Le questionnaire doit être clair et compréhensible par tous les individus cibles. Quant aux variables, celles-ci doivent être bien choisies afin d'aboutir à des résultats cohérentes. Pour mener à bien cette étape nous étions aidées par notre professeur de la Théorie de sondage Mme Mallek HELA qui nous a guidées tout le long de cette phase. Ensemble, nous avons élaboré un questionnaire de sept (7) questions.

### **2.3 Collecte de données**

Après l'élaboration du questionnaire, nous sommes passés au recueil de données. Nous avons collecté 100 réponses auprès des différents personnes après avoir publié le questionnaire sur les réseaux sociaux.

### **2.4 Description de la base de données**

Nous avons constitué avec les réponses collectées une table EXCEL comportant 100 observations et 7 variables

- Sexe : Sexe de l'individu
- Age : Age de l'individu
- Etat\_matrimonial : État matrimonial de l'individu
- Niveau\_etude : Niveau d'étude de l'individu
- CSP : Catégorie socioprofessionnelle de l'individu
- Budget : Budget en dinars consacré par l'individu à la friperie pendant les 3 derniers mois

---

## Enquête sur les dépenses en friperie

1. Sexe \*

- Femme
- Homme

2. Age \*

- [18,25[
- [25,40[
- [40,55[
- [55,∞[

3. Etat matrimonial \*

- Célibataire
- Marié
- Divorcé
- Veuf

4. Niveau d'étude \*

- Analphabète
- Primaire
- Secondaire
- Supérieur

5. Catégorie Socio-professionnelle (Chef de ménage) \*

- Sans emploi
- Ouvrier/Employé
- Agriculteur/Artisan/commerçant
- Cadre supérieur/Profession libérale

6. Quel budget avez-vous consacré pour la friperie pendant les 3 derniers mois ? \*

- [0,10[
- [10,30[
- [30,50[
- [50,∞[

7. Qu'est-ce que vous préférez acheter le plus ? \*

- Vêtements
- Accessoires
- Linge de maison
- Autre : -----

Fourni par



FIGURE 1 – Questionnaire

- Préférence : ce que l'individu préfère acheter le plus de la friperie.

On avait également calculé le budget moyen pour pouvoir faire le projet en la considérant comme variable d'intérêt.

## 2.5 Étude descriptive

Afin de connaître l'évolution des différentes variables dans notre échantillon, nous avons réalisé une étude descriptive sur celles-ci. Les résultats se présentent comme suit :

On a élaboré les camemberts des variables qualitatives

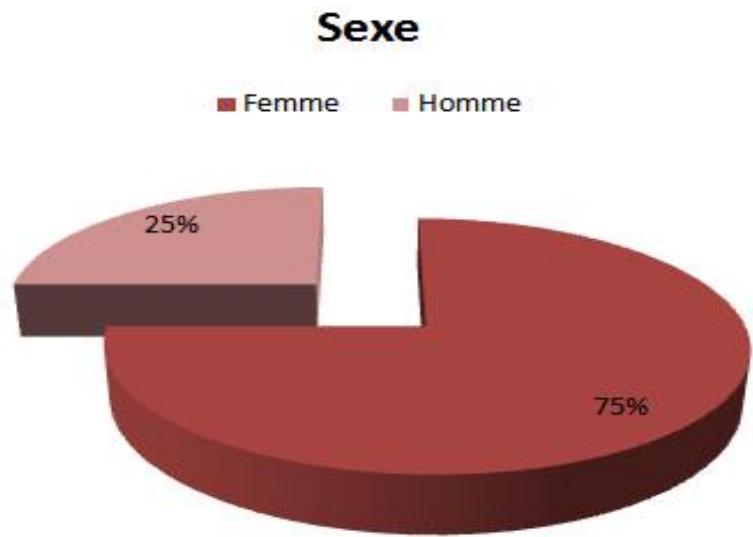


FIGURE 2 – Camembert de la variable sexe

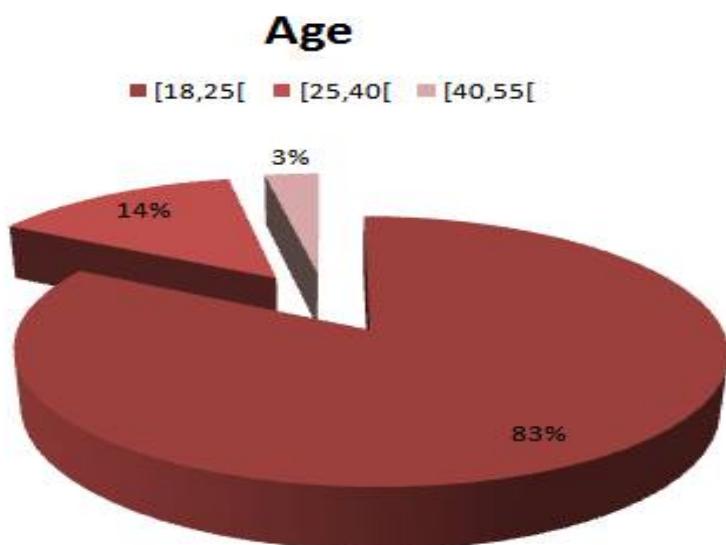


FIGURE 3 – Camembert de la variable age

---

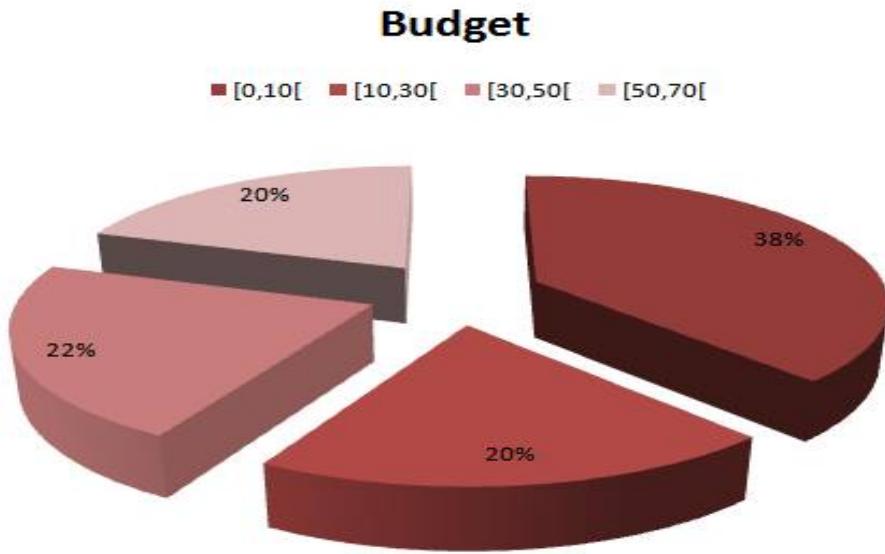


FIGURE 4 – Camembert de la variable budget

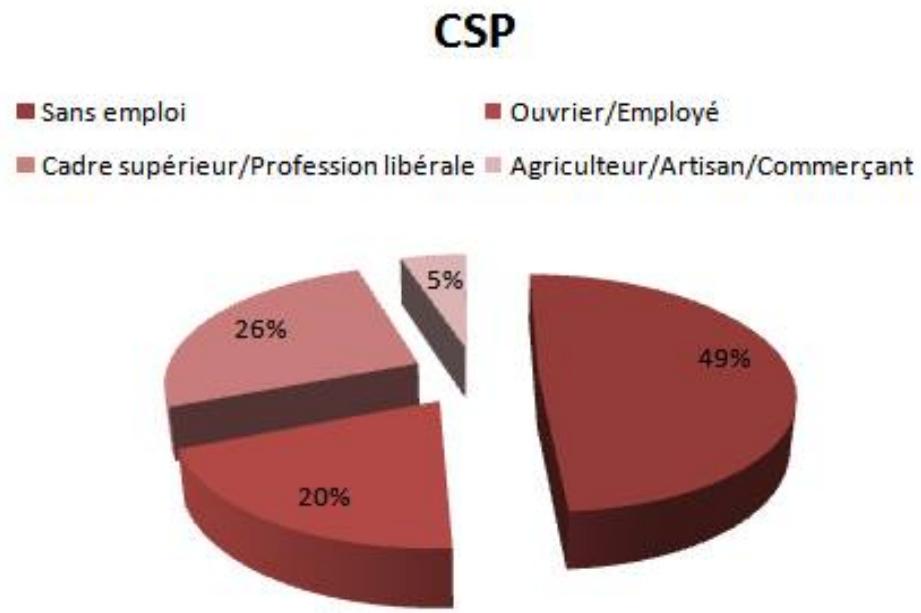


FIGURE 5 – Camembert de la variable catégorie socio-professionnelle



FIGURE 6 – Camembert de la variable état matrimonial

## Niveau d'étude

■ Supérieur ■ Secondaire ■ Primaire

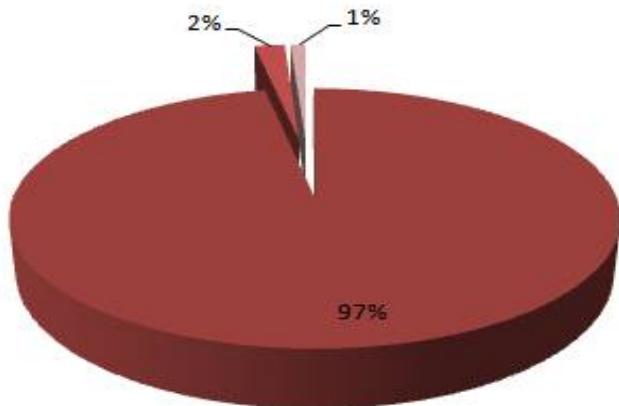


FIGURE 7 – Camembert de la variable niveau d'étude

## Préférence

■ Vêtements  
■ Accessoires  
■ Linge de maison  
■ Autres  
■ Vêtements, Accessoires  
■ Vêtements, Linge de maison  
■ Vêtements, Accessoires, Linge de maison

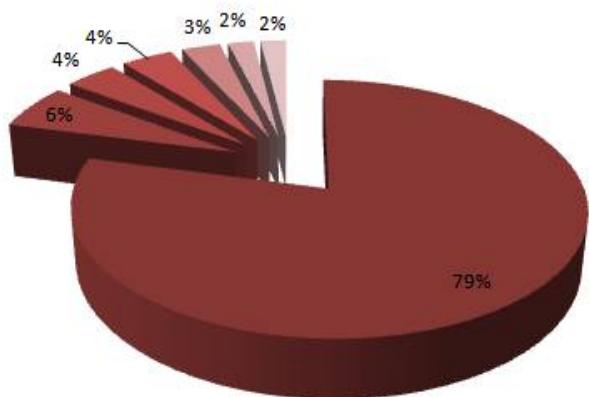


FIGURE 8 – Camembert de la variable préférence d'achat à la friperie

---

On a ensuite visualisé la variable d'interet VI qui est le budget moyen en fonction des variables qualitatives

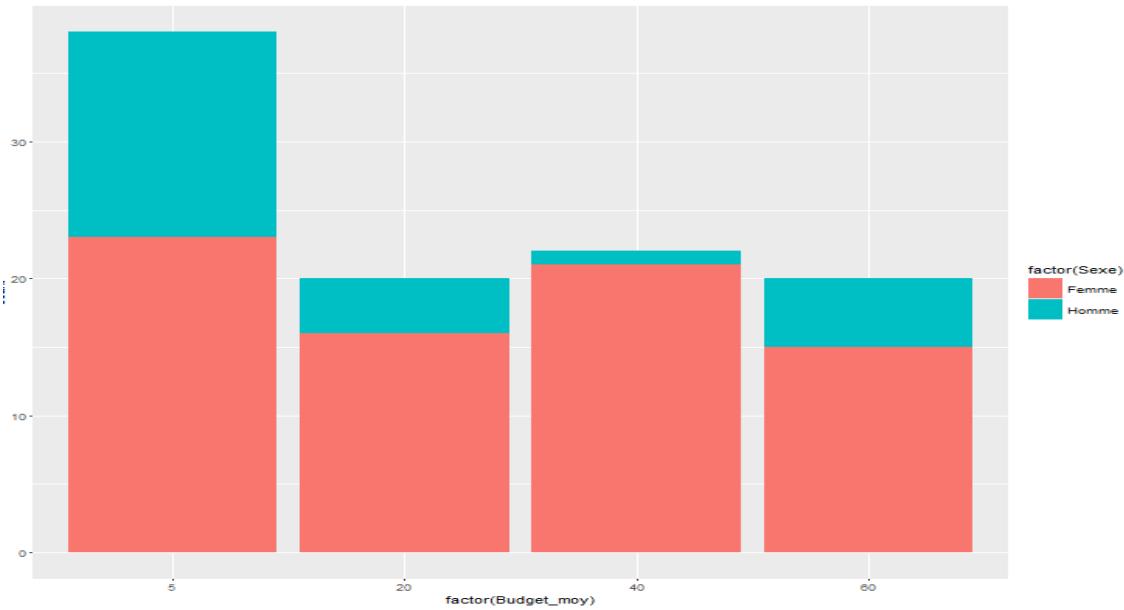


FIGURE 9 – Budget moyen en fonction de la variable sexe

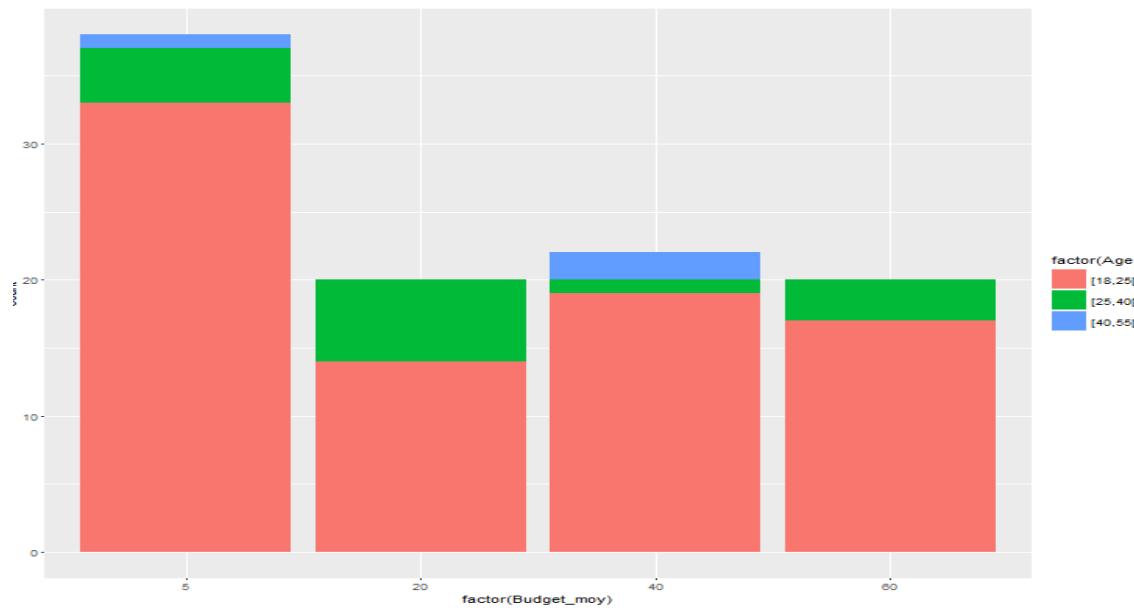


FIGURE 10 – Budget moyen en fonction de la variable age

---

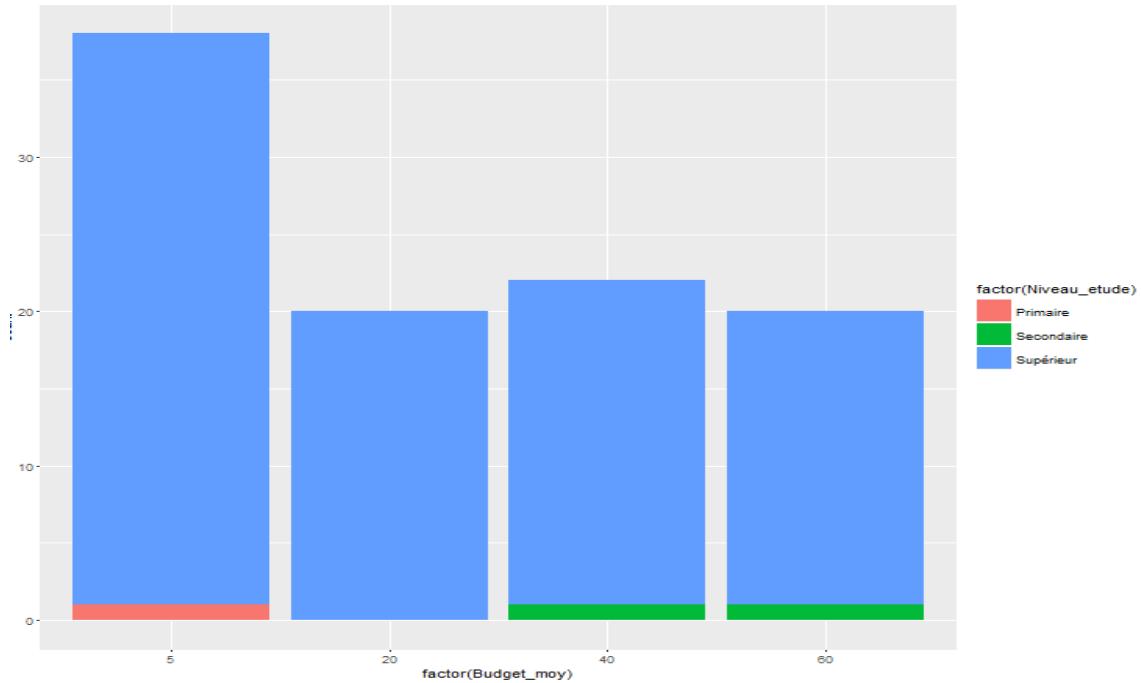


FIGURE 11 – Budget moyen en fonction de la variable niveau d'étude

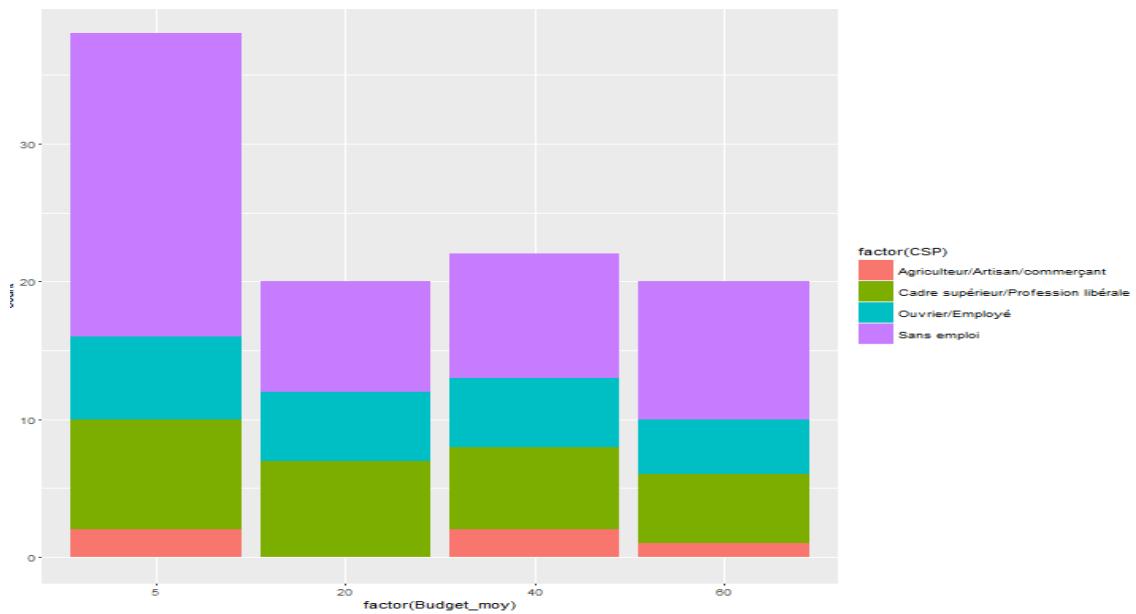


FIGURE 12 – Budget moyen en fonction de la variable catégorie socio professionnelle

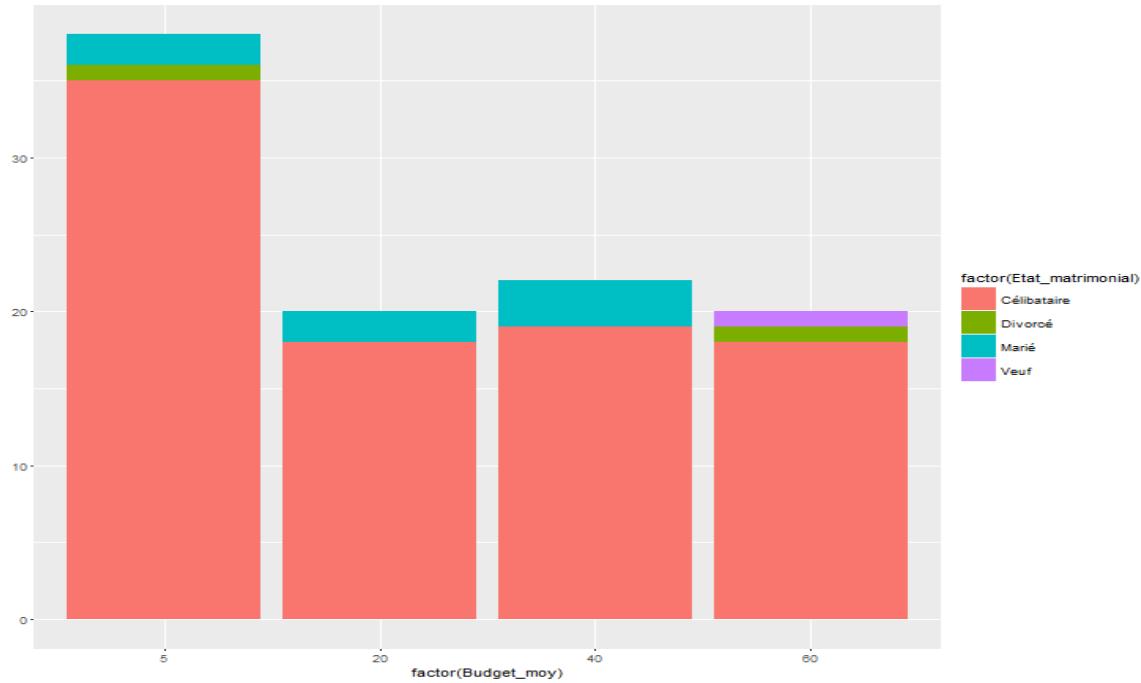


FIGURE 13 – Budget moyen en fonction de la variable niveau d'étude

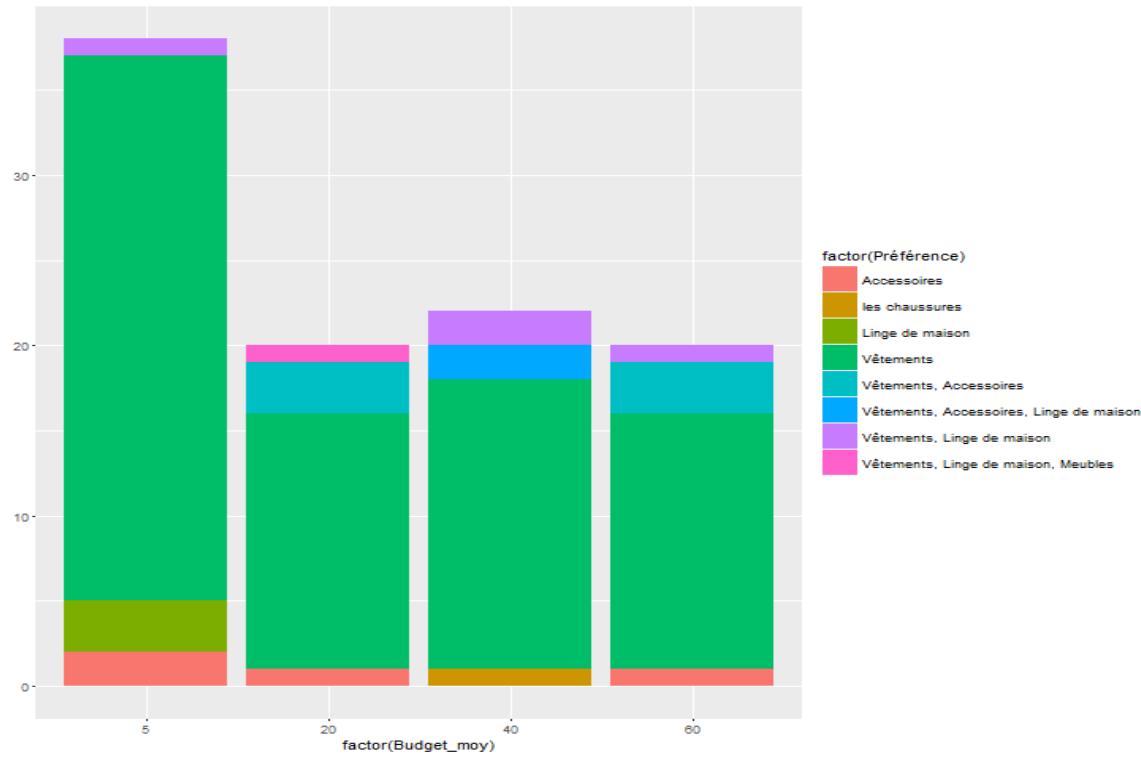


FIGURE 14 – Budget moyen en fonction de la variable préférence

Pour des raisons de besoin et de simplification, on a fait un codage des variables qualitatives pour les transformer en variables catégorielles pour pouvoir appliquer les méthodes de sondage stratifié et par grappes. On a établit ainsi la matrice de corrélation :

	Budget_moy	Agemoy	CSPm	Etat_matrimonialm	Niveau_etudem
Sexem					
Budget_moy	1.000000000	-0.002041668	0.02513647	0.10680320	0.006251858

FIGURE 15 – Corrélation des variables auxiliaires avec la VI

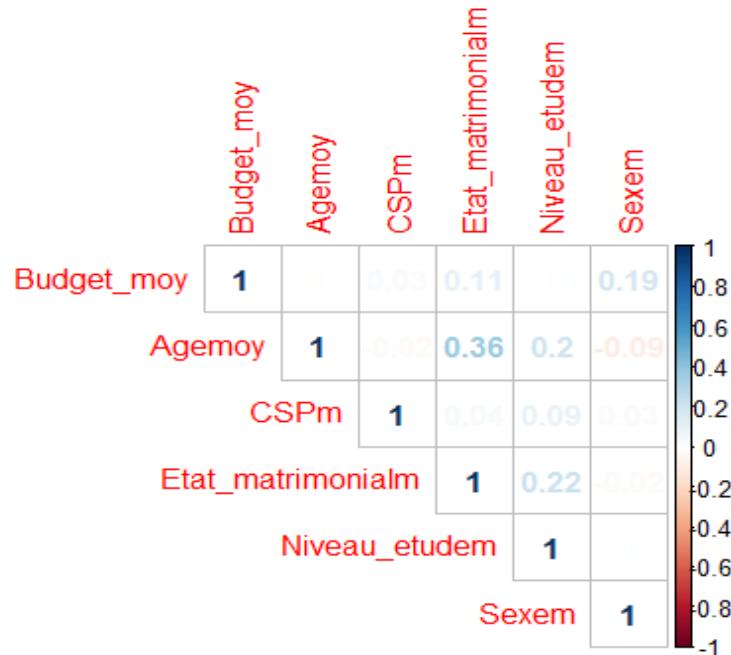


FIGURE 16 – Matrice de corrélation

### 3 Partie Analytique

Pour simuler la réalité, nous allons considérer notre base de 100 observations comme une base de sondage. Pour chaque méthode, nous tirerons de cette base k échantillons de  $n=20$  individus.

Ensuite nous effectuerons des opérations d'estimation (ponctuelle et par intervalle de confiance).

Les k intervalles de confiance exprimés en fonction des variances estimées correspondantes seront représentés sur un seul graphique.

A noter que : - L'unité utilisée est le dinar tunisien.

- Sur les 100 observations la variable d'intérêt a une moyenne = 26.7 dinars et une variance = 453.1414 .

- Pour les codes consultez l'ANNEXE.

Nous avons utilisé le logiciel RStudio Version 3.0.3 pour réaliser notre travail.

---

### 3.1 Sondage à probabilités égales

#### 3.1.1 Sans remise(PESR)

Cet échantillonnage consiste à sélectionner des individus au hasard selon une même probabilité et sans remise (i.e. un individu ne peut être sélectionné qu'une fois).

Voici les résultats obtenus avec cette méthode :

- Les 20 échantillons

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	E20
1	40	5	5	20	20	20	5	40	5	60	5	20	5	20	40	20	5	40	40	20
2	20	40	20	5	20	60	60	40	40	20	60	40	60	5	5	5	60	60	40	20
3	40	60	20	5	5	60	40	5	20	20	40	40	40	40	5	60	20	20	5	60
4	5	20	5	20	20	60	5	40	40	40	60	60	40	5	40	60	5	40	60	20
5	5	20	20	40	5	20	40	60	40	20	20	5	5	5	5	60	40	20	60	20
6	20	5	5	5	60	5	5	20	60	40	5	20	5	60	60	20	40	5	60	20
7	60	5	60	60	5	20	20	60	5	5	60	60	5	20	60	5	40	40	5	20
8	60	60	40	5	20	5	5	60	40	60	60	40	20	40	60	20	20	60	40	40
9	20	40	20	40	60	20	20	40	5	5	5	60	5	5	40	20	5	40	5	40
10	40	5	5	20	5	40	40	20	60	60	60	5	5	5	20	40	5	20	5	60
11	40	20	5	5	40	20	60	20	20	20	60	60	40	60	5	60	5	40	40	40
12	20	20	20	60	40	60	20	40	40	20	60	20	60	40	5	5	40	5	5	40
13	60	5	5	20	20	60	5	40	5	40	5	5	20	60	40	60	40	20	5	5
14	20	5	40	40	60	40	20	60	5	5	5	40	20	5	5	60	5	20	60	5
15	40	5	20	60	5	20	40	5	60	5	60	5	60	60	40	5	60	5	5	20
16	60	5	60	20	40	40	60	5	60	40	5	5	5	60	60	60	5	40	5	5
17	5	20	5	40	40	40	40	40	20	5	40	40	5	60	20	5	5	40	40	5
18	5	20	40	20	20	5	5	60	20	5	20	60	5	20	20	40	20	5	5	20
19	20	60	5	5	60	60	60	40	20	60	5	60	60	40	60	40	40	5	20	40
20	5	5	5	60	5	5	20	60	60	20	40	60	5	5	40	60	5	60	5	60

FIGURE 17 – Matrice d'échantillons PESR

- Les estimations

	ybar.pesr	var.pesr	varpesr.estim	ICinf	ICsup	long.IC
1	29.25	411.2500	16.45000	21.30051	37.19949	15.89897
2	21.25	399.6711	15.98684	13.41322	29.08678	15.67355
3	20.25	338.0921	13.52368	13.04219	27.45781	14.41563
4	27.50	432.8947	17.31579	19.34400	35.65600	16.31200
5	27.50	432.8947	17.31579	19.34400	35.65600	16.31200
6	33.00	458.9474	18.35789	24.60216	41.39784	16.79568
7	28.50	437.1053	17.48421	20.30443	36.69557	16.39114
8	37.75	377.5658	15.10263	30.13303	45.36697	15.23394
9	31.25	452.3026	18.09211	22.91318	39.58682	16.67365
10	27.50	432.8947	17.31579	19.34400	35.65600	16.31200
11	33.75	620.7237	24.82895	23.98358	43.51642	19.53283
12	35.25	509.1447	20.36579	26.40482	44.09518	17.69036
13	23.50	505.5263	20.22105	14.68631	32.31369	17.62739
14	30.75	550.7237	22.02895	21.55074	39.94926	18.39852
15	29.75	509.1447	20.36579	20.90482	38.59518	17.69036
16	34.25	519.1447	20.76579	25.31838	43.18162	17.86324
17	26.00	435.7895	17.43158	17.81678	34.18322	16.36645
18	26.50	342.3684	13.69474	19.24675	33.75325	14.50651
19	28.25	561.2500	22.45000	18.96324	37.53676	18.57352
20	28.00	337.8947	13.51579	20.79429	35.20571	14.41142

FIGURE 18 – Matrice des estimations PESR

- La moyenne = 28.9875
- Les intervalles de confiance IC(0.95)

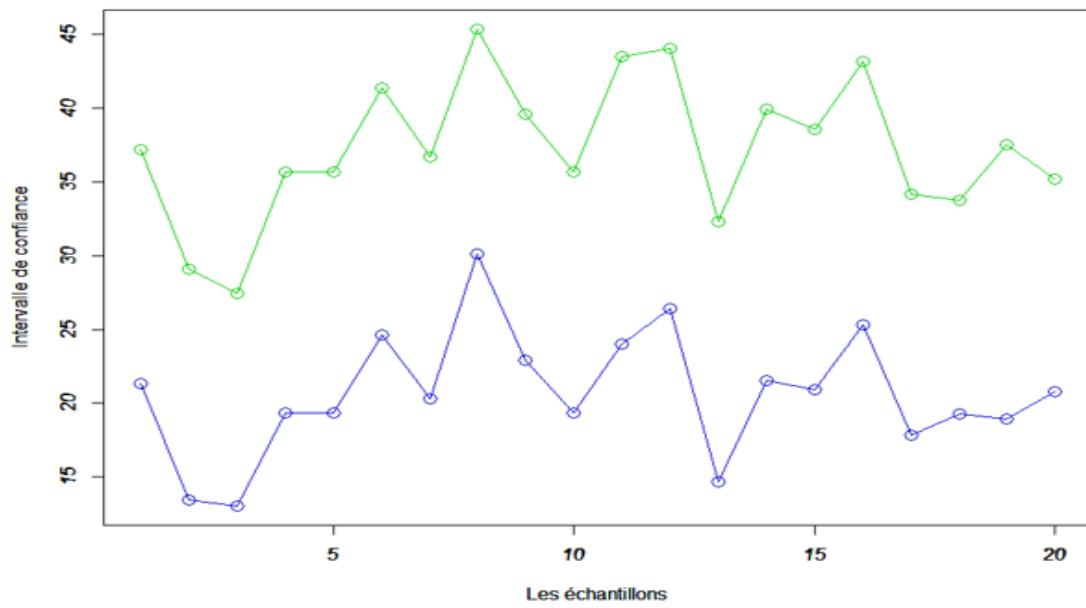


FIGURE 19 – Graphe des intervalles de confiance PESR

---

### 3.1.2 Avec remise (PEAR)

Il s'agit de répéter n fois et de façon indépendante la procédure de tirage d'un individu dans la population avec une même probabilité soit  $p=1/N$ .

Remarque : Avec cette méthode un même individu peut être tiré plusieurs fois. Cette possibilité de redondance d'informations n'est pas en faveur de la précision des estimations.

— Les 20 échantillons

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	E20
1	5	5	60	40	5	40	20	40	20	40	5	60	5	20	60	40	20	60	60	5
2	60	5	5	5	40	5	5	5	60	60	60	40	20	5	20	5	60	60	20	5
3	60	40	60	5	20	20	40	40	40	40	5	5	5	20	5	5	60	5	60	20
4	40	60	40	20	40	5	40	5	40	5	20	60	40	60	40	60	5	5	60	20
5	60	60	60	20	40	5	40	40	5	5	5	60	20	20	60	5	5	60	20	
6	5	60	20	40	40	40	5	20	40	40	5	5	60	40	60	5	20	60	5	20
7	40	40	60	40	40	60	20	60	40	5	60	5	5	5	20	60	5	5	60	5
8	20	5	5	40	60	5	20	5	60	5	60	20	40	5	5	40	5	60	60	5
9	60	60	5	20	5	20	60	40	60	20	5	5	60	5	60	5	5	60	40	5
10	5	5	40	60	5	60	60	20	40	40	40	40	5	5	5	5	40	5	20	20
11	5	20	5	60	40	5	60	60	5	5	60	5	20	20	5	20	5	5	5	5
12	40	40	5	40	5	60	40	20	5	20	5	40	40	40	5	20	40	40	5	60
13	20	40	5	60	20	5	5	5	5	40	60	5	5	5	5	20	40	20	5	
14	5	20	5	60	5	40	5	5	60	60	5	20	60	5	5	5	40	40	5	60
15	5	5	60	5	40	20	20	40	40	5	40	60	20	5	20	5	5	5	60	5
16	60	60	5	60	5	5	5	5	60	5	5	60	5	20	5	60	20	40	60	40
17	5	40	20	5	60	20	40	40	40	5	60	40	60	40	40	5	20	5	20	20
18	5	5	20	20	20	5	60	40	40	60	60	5	40	40	40	5	5	20	40	60
19	5	5	20	60	5	40	20	20	60	5	5	5	20	20	60	5	40	40	40	40
20	20	5	5	40	5	40	60	20	60	5	40	5	60	20	60	60	5	20	40	60

FIGURE 20 – Matrice d'échantillons PEAR

— Les estimations

	ybar.pear	var.pear	varpear.estim	ICinf	ICsup	long.IC
1	26.25	549.6711	21.98684	17.05953	35.44047	18.38093
2	29.00	535.7895	21.43158	19.92633	38.07367	18.14735
3	25.25	540.7237	21.62895	16.13464	34.36536	18.23072
4	35.00	442.1053	17.68421	26.75769	43.24231	16.48462
5	25.00	384.2105	15.36842	17.31629	32.68371	15.36741
6	25.00	426.3158	17.05263	16.90621	33.09379	16.18757
7	31.25	452.3026	18.09211	22.91318	39.58682	16.67365
8	26.50	342.3684	13.69474	19.24675	33.75325	14.50651
9	39.00	425.2632	17.01053	30.91621	47.08379	16.16758
10	21.75	463.8816	18.55526	13.30714	30.19286	16.88572
11	29.25	605.9868	24.23947	19.60021	38.89979	19.29957
12	30.00	578.9474	23.15789	20.56796	39.43204	18.86408
13	29.50	481.3158	19.25263	20.89995	38.10005	17.20011
14	20.00	263.1579	10.52632	13.64092	26.35908	12.71816
15	29.00	577.8947	23.11579	19.57654	38.42346	18.84692
16	21.00	520.0000	20.80000	12.06102	29.93898	17.87795
17	21.25	357.5658	14.30263	13.83751	28.66249	14.82498
18	29.00	535.7895	21.43158	19.92633	38.07367	18.14735
19	37.00	501.0526	20.04211	28.22539	45.77461	17.54922
20	24.00	456.8421	18.27368	15.62144	32.37856	16.75711

FIGURE 21 – Matrice des estimations PEAR

- La moyenne = 27.38125
- Les intervalles de confiance IC(0.95)

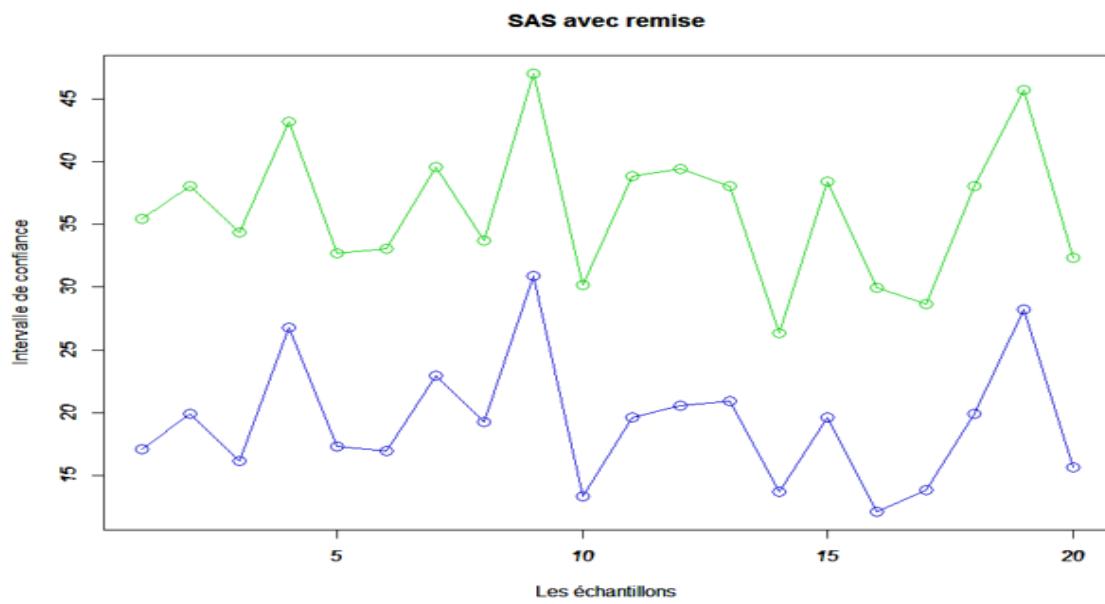


FIGURE 22 – Graphe des intervalles de confiance PEAR

## 3.2 Sondage à probabilités inégales

Contrairement à probabilités égales, le sondage à probabilités inégales suppose des probabilités d'inclusion différentes pour les individus. Il nécessite la prise en compte d'une variable auxiliaire positivement corrélée avec la variable d'intérêt. Les probabilités d'inclusion sont fonction de la variable auxiliaire afin de minimiser la variance et ainsi augmenter la précision. Nous avons choisi comme variable auxiliaire le temps restant à l'individu après le travail personnel, soit VAX=24travail\_perso. 2

### 3.2.1 Avec remise PIAR

- Les 20 échantillons

	Ech1	Ech2	Ech3	Ech4	Ech5	Ech6	Ech7	Ech8	Ech9	Ech10	Ech11	Ech12	Ech13	Ech14	Ech15	Ech16	Ech17	Ech18	Ech19	Ech20
1	60	60	20	60	20	60	40	60	40	60	60	60	40	60	60	60	40	60	40	60
2	20	60	60	20	40	40	40	60	20	60	60	60	20	60	60	60	60	40	60	40
3	40	20	20	5	60	60	40	40	20	20	40	40	60	5	40	20	60	60	40	20
4	40	60	60	60	40	20	40	60	60	60	60	60	40	5	60	40	5	20	40	5
5	60	40	40	20	60	40	60	60	60	60	60	60	5	40	40	40	60	60	60	60
6	60	20	40	60	20	40	60	5	40	60	40	20	40	60	60	40	40	40	60	60
7	60	40	5	40	60	40	40	60	20	40	60	20	40	60	60	60	40	40	40	40
8	40	60	60	40	60	40	5	60	60	60	40	40	40	60	40	60	60	40	40	20
9	40	60	40	20	20	5	60	60	60	20	40	60	60	5	60	40	60	60	40	40
10	60	20	40	60	60	60	60	40	60	60	60	40	20	40	40	60	40	60	40	60
11	40	40	20	60	40	20	20	40	60	40	5	60	40	60	60	20	60	40	5	60
12	60	40	40	60	40	60	40	20	40	60	60	5	5	40	40	60	20	60	20	40
13	60	40	40	60	60	20	60	5	40	20	60	20	60	60	60	40	40	60	40	40
14	40	20	60	20	40	60	60	60	40	40	60	60	40	20	60	40	40	20	20	5
15	60	60	5	60	40	60	40	60	40	60	60	60	60	60	60	40	60	40	60	60
16	60	40	60	60	60	40	20	60	60	60	40	60	60	60	60	40	60	20	20	20
17	60	60	60	60	60	20	40	40	40	40	60	60	60	40	60	60	40	60	60	40
18	60	60	40	40	60	40	40	60	60	60	60	40	60	40	60	60	20	5	60	40
19	60	40	60	60	60	60	60	60	40	60	60	20	40	40	60	40	60	60	60	60
20	40	60	20	60	40	60	5	60	60	60	60	40	5	5	40	40	5	60	60	5

FIGURE 23 – Matrice d'échantillons PIAR

- Les estimations

	<b>Ybarpiar</b>	<b>VarYbarpiar_Est</b>	<b>BInf_IC</b>	<b>BSup_IC</b>	<b>Longueur</b>
Ech1	<b>51.00</b>	<b>146.31579</b>	<b>49.10677</b>	<b>52.89323</b>	<b>3.786457</b>
Ech2	<b>45.00</b>	<b>247.36842</b>	<b>43.10677</b>	<b>46.89323</b>	<b>3.786457</b>
Ech3	<b>39.50</b>	<b>360.26316</b>	<b>37.60677</b>	<b>41.39323</b>	<b>3.786457</b>
Ech4	<b>46.25</b>	<b>360.19737</b>	<b>44.35677</b>	<b>48.14323</b>	<b>3.786457</b>
Ech5	<b>47.00</b>	<b>222.10526</b>	<b>45.10677</b>	<b>48.89323</b>	<b>3.786457</b>
Ech6	<b>42.25</b>	<b>311.77632</b>	<b>40.35677</b>	<b>44.14323</b>	<b>3.786457</b>
Ech7	<b>41.50</b>	<b>316.05263</b>	<b>39.60677</b>	<b>43.39323</b>	<b>3.786457</b>
Ech8	<b>46.50</b>	<b>337.10526</b>	<b>44.60677</b>	<b>48.39323</b>	<b>3.786457</b>
Ech9	<b>50.00</b>	<b>147.36842</b>	<b>48.10677</b>	<b>51.89323</b>	<b>3.786457</b>
Ech10	<b>47.00</b>	<b>306.31579</b>	<b>45.10677</b>	<b>48.89323</b>	<b>3.786457</b>
Ech11	<b>51.25</b>	<b>204.93421</b>	<b>49.35677</b>	<b>53.14323</b>	<b>3.786457</b>
Ech12	<b>46.25</b>	<b>318.09211</b>	<b>44.35677</b>	<b>48.14323</b>	<b>3.786457</b>
Ech13	<b>40.75</b>	<b>403.35526</b>	<b>38.85677</b>	<b>42.64323</b>	<b>3.786457</b>
Ech14	<b>38.00</b>	<b>443.15789</b>	<b>36.10677</b>	<b>39.89323</b>	<b>3.786457</b>
Ech15	<b>54.00</b>	<b>88.42105</b>	<b>52.10677</b>	<b>55.89323</b>	<b>3.786457</b>
Ech16	<b>46.00</b>	<b>172.63158</b>	<b>44.10677</b>	<b>47.89323</b>	<b>3.786457</b>
Ech17	<b>43.50</b>	<b>347.63158</b>	<b>41.60677</b>	<b>45.39323</b>	<b>3.786457</b>
Ech18	<b>44.25</b>	<b>298.09211</b>	<b>42.35677</b>	<b>46.14323</b>	<b>3.786457</b>
Ech19	<b>44.25</b>	<b>298.09211</b>	<b>42.35677</b>	<b>46.14323</b>	<b>3.786457</b>
Ech20	<b>38.75</b>	<b>402.30263</b>	<b>36.85677</b>	<b>40.64323</b>	<b>3.786457</b>

FIGURE 24 – Matrice des estimations PIAR

- La moyenne = 27.38125
- Les intervalles de confiance IC(0.95)

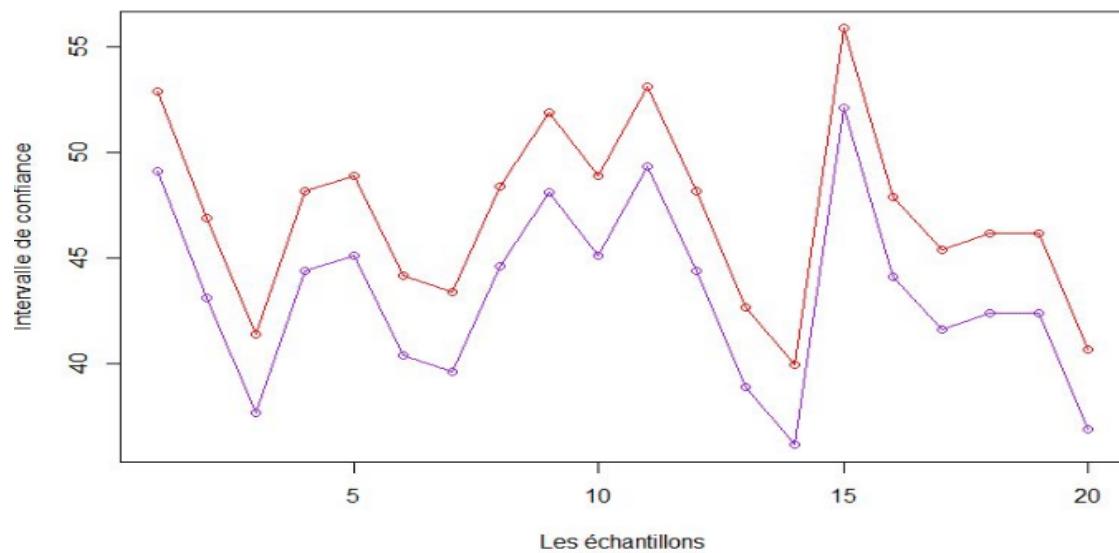


FIGURE 25 – Graphe des intervalles de confiance PIAR

---

### 3.2.2 Sans remise PISR

— Les 20 échantillons

	Ech1	Ech2	Ech3	Ech4	Ech5	Ech6	Ech7	Ech8	Ech9	Ech10	Ech11	Ech12	Ech13	Ech14	Ech15	Ech16	Ech17	Ech18	Ech19	Ech20
1	60	60	20	60	20	60	40	60	40	60	60	60	40	60	60	60	40	60	40	60
2	20	60	60	20	40	40	40	40	60	20	60	60	60	20	60	60	60	40	60	40
3	40	20	20	5	60	60	40	40	20	20	40	40	60	5	40	20	60	60	40	20
4	40	60	60	60	40	20	40	60	60	60	60	60	40	5	60	40	5	20	40	5
5	60	40	40	20	60	40	60	60	60	60	60	60	5	40	40	40	60	60	60	60
6	60	20	40	60	20	40	60	5	40	60	40	20	40	60	60	40	40	40	60	60
7	60	40	5	40	60	40	40	40	60	20	40	60	20	40	60	60	40	40	40	40
8	40	60	60	40	60	40	5	60	60	60	40	40	40	60	40	60	60	40	40	20
9	40	60	40	20	20	5	60	60	60	20	40	60	60	5	60	40	60	40	60	40
10	60	20	40	60	60	60	40	60	60	60	60	40	20	40	40	60	40	60	40	60
11	40	40	20	60	40	20	20	40	60	40	5	60	40	60	60	20	60	40	5	60
12	60	40	40	60	40	60	40	20	40	60	60	5	5	40	40	60	20	60	20	40
13	60	40	40	60	60	20	60	5	40	20	60	20	60	60	60	40	40	60	40	40
14	40	20	60	20	40	60	60	60	40	40	60	60	40	20	60	40	40	20	20	5
15	60	60	5	60	40	60	40	60	40	60	60	60	60	60	60	40	60	40	60	60
16	60	40	60	60	60	40	20	60	60	60	40	60	60	60	60	40	60	20	20	20
17	60	60	60	60	60	20	40	40	40	40	60	60	60	40	60	60	40	60	60	40
18	60	60	40	40	60	40	40	60	60	60	60	40	60	40	60	60	20	5	60	40
19	60	40	60	60	60	60	60	60	40	60	60	20	40	40	60	40	60	60	60	60
20	40	60	20	60	40	60	5	60	60	60	60	40	5	5	40	40	5	60	60	5

FIGURE 26 – Matrice d'échantillons PISR

— Les estimations

	<b>Ybarpisr</b>	<b>VarYbarpisr_Est</b>	<b>BInf_IC</b>	<b>BSup_IC</b>	<b>Longueur</b>
Ech1	27.69368	23.29575	27.57715	27.81020	0.2330497
Ech2	27.66548	24.56400	27.54895	27.78200	0.2330497
Ech3	27.69361	24.29700	27.57709	27.81014	0.2330497
Ech4	27.18476	24.36375	27.06823	27.30128	0.2330497
Ech5	27.70512	23.82975	27.58860	27.82165	0.2330497
Ech6	27.70190	23.82975	27.58537	27.81842	0.2330497
Ech7	27.17287	24.89775	27.05635	27.28940	0.2330497
Ech8	27.67456	24.56400	27.55804	27.79109	0.2330497
Ech9	27.19024	24.03000	27.07372	27.30677	0.2330497
Ech10	27.17174	25.09800	27.05522	27.28827	0.2330497
Ech11	27.68267	24.23025	27.56614	27.79919	0.2330497
Ech12	27.15555	24.29700	27.03902	27.27207	0.2330497
Ech13	27.15959	23.49600	27.04306	27.27611	0.2330497
Ech14	27.67664	23.89650	27.56011	27.79316	0.2330497
Ech15	27.67664	23.89650	27.56011	27.79316	0.2330497
Ech16	27.70958	24.56400	27.59306	27.82611	0.2330497
Ech17	27.14178	23.29575	27.02526	27.25831	0.2330497
Ech18	27.70957	24.56400	27.59305	27.82610	0.2330497
Ech19	27.19502	23.76300	27.07850	27.31155	0.2330497
Ech20	27.15852	24.09675	27.04200	27.27505	0.2330497

FIGURE 27 – Matrice des estimations PISR

- La moyenne = 27.38125
- Les intervalles de confiance IC(0.95)

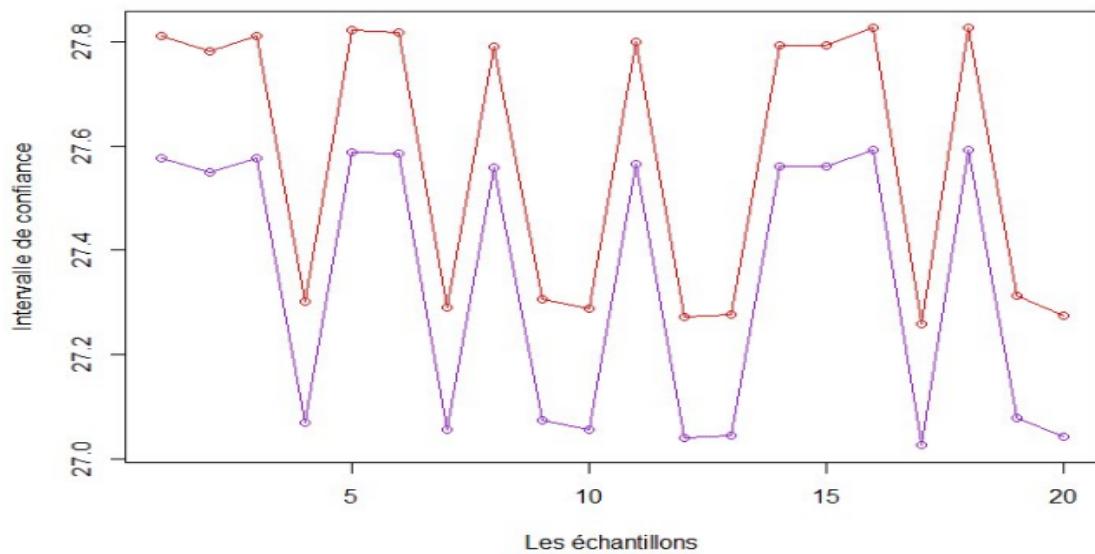


FIGURE 28 – Graphe des intervalles de confiance PISR

### 3.3 Startification

La stratification est une méthode permettant d'introduire de l'information auxiliaire dans le plan de sondage. Elle consiste à partitionner la base de sondage en M sous-ensembles et à procéder, de façon indépendante à un sondage dans chaque sous-population (ou strate).

Nous avons choisi la variable Genre comme variable auxiliaire, la population a été donc partitionnée selon le genre.

#### 3.3.1 Strates à tailles égales

Nous allons tirer nh=10 individus dans chacune des strates

— Les échantillons

	Ech1	Ech2	Ech3	Ech4	Ech5	Ech6	Ech7	Ech8	Ech9	Ech10	Ech11	Ech12	Ech13	Ech14	Ech15	Ech16	Ech17	Ech18	Ech19	Ech20
1	20	60	5	5	5	20	20	60	5	20	20	20	20	60	60	60	60	60	60	60
2	5	60	20	20	20	20	5	5	5	20	40	5	20	20	5	60	5	60	20	60
3	5	5	5	5	5	5	5	40	40	5	5	5	5	40	20	5	20	5	5	5
4	5	5	5	40	5	5	5	5	5	5	5	40	5	5	5	5	40	5	5	20
5	40	5	5	5	60	60	5	5	5	5	5	5	5	60	5	5	5	5	5	5
6	5	5	5	5	5	5	60	5	60	20	5	60	60	5	5	5	5	20	5	40
7	5	20	60	60	60	60	20	60	5	20	60	20	20	5	20	40	60	60	5	5
8	60	20	5	5	20	20	5	5	5	20	20	20	5	5	5	5	20	20	5	5
9	20	5	60	5	5	5	60	5	5	5	5	5	5	5	5	5	5	5	5	5
10	60	60	20	5	60	60	5	5	60	60	5	5	60	5	5	60	5	60	60	5
11	40	5	60	5	40	60	20	40	5	60	60	40	60	60	60	60	40	40	5	40
12	20	5	40	60	60	60	40	40	5	60	60	40	20	40	20	5	5	5	60	40
13	40	40	5	60	5	5	40	5	5	5	20	60	5	5	40	60	20	5	5	5
14	40	40	40	40	40	40	20	5	5	40	20	40	5	40	20	60	60	40	5	20
15	5	60	60	40	40	20	5	40	5	5	5	40	40	5	20	60	40	40	60	5
16	20	5	40	40	20	5	20	20	40	20	40	5	5	60	40	60	60	20	5	5
17	60	5	20	20	60	40	40	5	60	5	20	40	60	20	40	60	60	5	40	40
18	20	60	5	40	5	40	20	5	60	40	5	40	20	40	60	5	40	5	5	60
19	40	40	20	5	5	5	40	60	5	5	20	60	5	5	60	40	5	5	20	5
20	60	20	5	40	40	20	5	5	40	60	20	40	60	5	5	5	5	40	20	40

FIGURE 29 – Matrice d'échantillons issus de la stratification à taille égale

— Les estimations

---

	<b>ybar.strate</b>	<b>var.estim.strate</b>	<b>ICinf</b>	<b>ICsup</b>	<b>IClong</b>
<b>1</b>	<b>31.500</b>	<b>17.77292</b>	<b>23.23704</b>	<b>39.76296</b>	<b>16.52591</b>
<b>2</b>	<b>27.125</b>	<b>27.62604</b>	<b>16.82315</b>	<b>37.42685</b>	<b>20.60371</b>
<b>3</b>	<b>26.875</b>	<b>24.49271</b>	<b>17.17494</b>	<b>36.57506</b>	<b>19.40012</b>
<b>4</b>	<b>30.125</b>	<b>19.82187</b>	<b>21.39873</b>	<b>38.85127</b>	<b>17.45253</b>
<b>5</b>	<b>29.750</b>	<b>24.87708</b>	<b>19.97412</b>	<b>39.52588</b>	<b>19.55176</b>
<b>6</b>	<b>28.625</b>	<b>24.82604</b>	<b>18.85916</b>	<b>38.39084</b>	<b>19.53169</b>
<b>7</b>	<b>22.125</b>	<b>10.89687</b>	<b>15.65496</b>	<b>28.59504</b>	<b>12.94008</b>
<b>8</b>	<b>23.125</b>	<b>23.54688</b>	<b>13.61408</b>	<b>32.63592</b>	<b>19.02185</b>
<b>9</b>	<b>22.125</b>	<b>30.64688</b>	<b>11.27451</b>	<b>32.97549</b>	<b>21.70097</b>
<b>10</b>	<b>26.625</b>	<b>30.87604</b>	<b>15.73402</b>	<b>37.51598</b>	<b>21.78196</b>
<b>11</b>	<b>23.375</b>	<b>22.87604</b>	<b>14.00053</b>	<b>32.74947</b>	<b>18.74893</b>
<b>12</b>	<b>32.000</b>	<b>11.59792</b>	<b>25.32508</b>	<b>38.67492</b>	<b>13.34984</b>
<b>13</b>	<b>30.250</b>	<b>32.41875</b>	<b>19.09026</b>	<b>41.40974</b>	<b>22.31949</b>
<b>14</b>	<b>19.875</b>	<b>22.07187</b>	<b>10.66678</b>	<b>29.08322</b>	<b>18.41644</b>
<b>15</b>	<b>30.000</b>	<b>25.67500</b>	<b>20.06858</b>	<b>39.93142</b>	<b>19.86284</b>
<b>16</b>	<b>37.750</b>	<b>35.24375</b>	<b>26.11418</b>	<b>49.38582</b>	<b>23.27165</b>
<b>17</b>	<b>33.000</b>	<b>23.50833</b>	<b>23.49686</b>	<b>42.50314</b>	<b>19.00627</b>
<b>18</b>	<b>26.250</b>	<b>24.72708</b>	<b>16.50364</b>	<b>35.99636</b>	<b>19.49272</b>
<b>19</b>	<b>21.250</b>	<b>27.21875</b>	<b>11.02437</b>	<b>31.47563</b>	<b>20.45126</b>
<b>20</b>	<b>25.875</b>	<b>20.00104</b>	<b>17.10939</b>	<b>34.64061</b>	<b>17.53123</b>

FIGURE 30 – Matrice des estimations par stratification à taille égale

- La moyenne = 27.38125
- Les intervalles de confiance

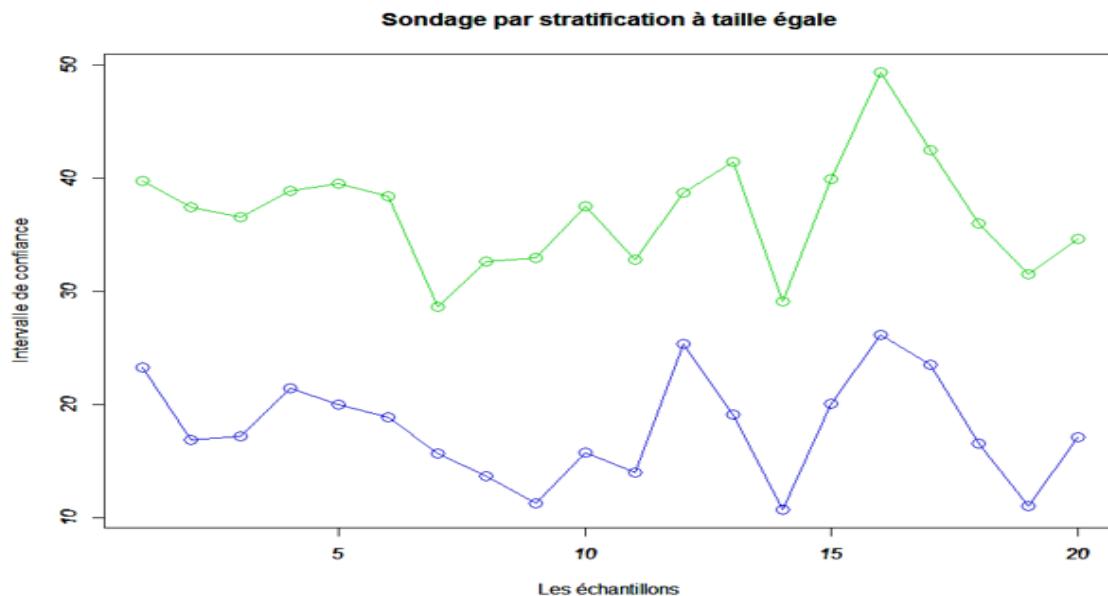


FIGURE 31 – Graphe des intervalles de confiance par la stratification à taille égale

### 3.3.2 Strates à allocations proportionnelles

— Les échantillons

	Ech1	Ech2	Ech3	Ech4	Ech5	Ech6	Ech7	Ech8	Ech9	Ech10	Ech11	Ech12	Ech13	Ech14	Ech15	Ech16	Ech17	Ech18	Ech19	Ech20
1	5	5	5	5	5	60	5	20	5	5	5	60	20	20	5	5	40	5	5	5
2	5	20	60	5	20	5	5	20	20	5	5	5	5	20	20	5	5	40	5	5
3	5	20	5	5	60	5	40	60	5	5	5	20	60	60	20	5	5	5	5	5
4	20	5	5	5	5	60	5	60	20	5	5	60	5	5	20	5	5	5	5	5
5	5	60	5	5	60	5	60	5	60	60	5	60	5	60	5	40	5	5	5	5
6	5	20	5	40	20	40	5	5	5	60	20	60	40	5	20	60	60	40	40	40
7	20	20	40	5	5	60	5	60	5	60	5	40	5	40	40	60	40	5	20	20
8	60	5	60	40	5	60	20	60	20	20	40	60	5	5	40	40	40	20	60	5
9	40	40	5	60	40	5	20	40	5	40	40	5	5	60	20	40	20	20	5	20
10	40	40	5	60	60	5	20	5	40	5	5	5	60	40	40	20	20	20	60	20
11	40	60	5	20	5	40	5	40	20	5	5	40	40	40	60	60	60	40	5	20
12	20	20	5	20	60	5	5	20	40	20	5	20	20	40	40	20	60	20	60	60
13	5	60	40	60	60	40	40	5	60	20	5	60	20	60	60	60	40	60	20	60
14	5	40	40	60	20	5	5	60	5	60	60	40	20	40	40	60	40	40	60	60
15	5	5	20	5	40	60	5	20	60	20	5	60	60	60	5	40	60	20	20	60
16	20	20	40	20	40	60	60	20	40	40	5	40	40	60	40	60	5	5	60	40
17	60	20	40	40	40	60	20	60	20	40	20	60	5	40	60	60	40	40	20	40
18	40	5	20	40	40	40	5	40	60	40	5	40	60	60	5	40	40	20	5	60
19	40	5	20	60	20	5	60	20	20	5	40	40	20	40	40	40	5	60	60	5
20	40	40	5	5	5	5	20	5	40	60	20	5	20	40	60	5	60	5	40	60

FIGURE 32 – Matrice d'échantillons à allocation proportionnelle

— Les estimations

---

	ybar.prop	var.estim.ybarprop	ICinf	ICsup	IClong
1	24.00	14.955556	16.420207	31.57979	15.15959
2	25.50	16.988889	17.421354	33.57865	16.15729
3	21.50	12.922222	14.454291	28.54571	14.09142
4	28.00	20.377778	19.152217	36.84778	17.69557
5	30.50	17.933333	22.199838	38.80016	16.60032
6	28.50	25.200000	18.660878	38.33912	19.67824
7	23.25	20.363889	14.405232	32.09477	17.68954
8	28.50	18.200000	20.138354	36.86165	16.72329
9	30.25	16.497222	22.289113	38.21089	15.92177
10	26.00	17.622222	17.772149	34.22785	16.45570
11	15.25	13.036111	8.173311	22.32669	14.15338
12	39.00	16.088889	31.138252	46.86175	15.72350
13	25.75	15.897222	17.935221	33.56478	15.62956
14	39.75	8.702778	33.967908	45.53209	11.56418
15	32.00	15.200000	24.358513	39.64149	15.28297
16	36.25	16.586111	28.267694	44.23231	15.96461
17	32.50	17.400000	24.324192	40.67581	16.35162
18	23.75	14.136111	16.380788	31.11921	14.73842
19	28.00	22.155556	18.774341	37.22566	18.45132
20	29.75	13.408333	22.572991	36.92701	14.35402

---

FIGURE 33 – Matrice des estimations par stratification à allocation proportionnelle

- La moyenne = 28.4
- Les intervalles de confiance

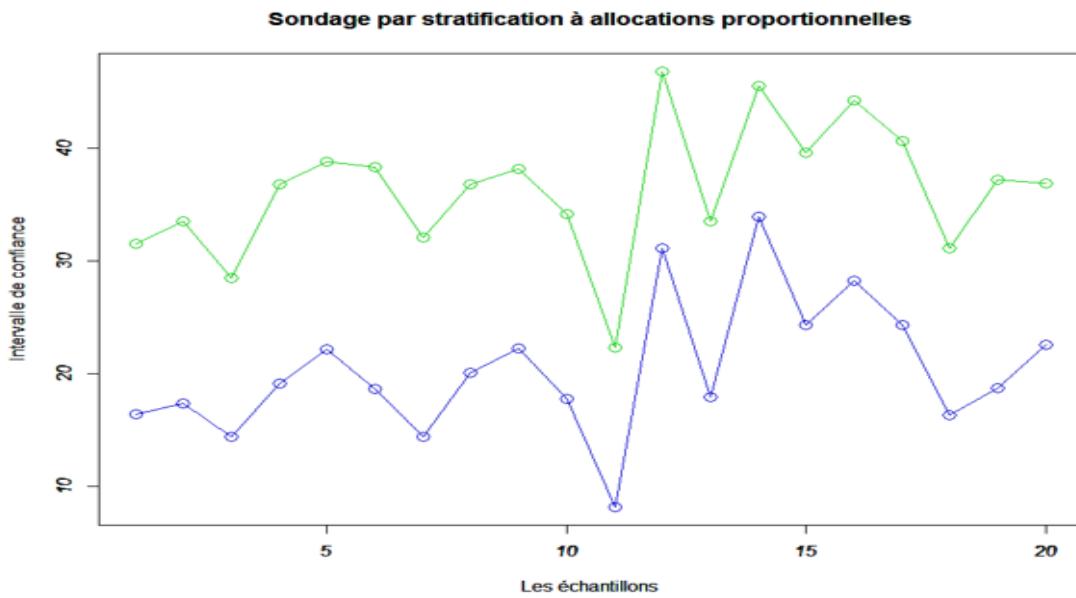


FIGURE 34 – Graphe des IC0(0.95) par stratification à allocations proportionnelles

### 3.3.3 Stratification à allocation optimale

Cette méthode permet d'optimiser la précision. Les nh seront à la fois proportionnelles à Nh et à la dispersion de la variable d'intérêt au sein des strates.

— Les échantillons

	Ech1	Ech2	Ech3	Ech4	Ech5	Ech6	Ech7	Ech8	Ech9	Ech10	Ech11	Ech12	Ech13	Ech14	Ech15	Ech16	Ech17	Ech18	Ech19	Ech20
1	5	5	5	5	5	60	20	5	5	20	5	5	20	5	60	5	5	20	5	5
2	5	5	5	5	20	5	60	5	5	5	60	5	40	20	5	20	60	5	60	5
3	5	60	5	60	5	5	5	5	20	5	60	5	5	5	60	5	5	5	5	60
4	20	5	5	5	5	5	5	5	40	5	60	60	20	5	5	5	5	20	5	60
5	5	5	5	5	60	5	5	5	5	20	5	5	5	5	60	60	5	5	20	
6	20	40	20	40	60	20	40	40	60	5	40	20	40	40	5	40	5	20	40	60
7	40	40	40	40	60	40	5	60	60	40	40	40	60	20	40	20	40	5	40	20
8	60	5	5	5	20	5	60	40	20	20	40	40	5	20	40	5	5	20	5	40
9	60	40	40	5	60	5	5	40	40	5	20	40	20	40	40	40	5	40	5	40
10	60	20	5	5	20	60	40	20	60	20	60	60	20	60	5	40	60	40	40	20
11	5	60	20	5	5	5	60	20	40	40	5	40	20	20	60	5	5	5	20	40
12	60	40	60	20	20	20	5	20	5	5	40	40	20	5	20	5	40	20	60	60
13	5	60	40	40	60	40	60	60	5	5	60	60	5	20	40	40	5	60	5	
14	40	60	5	40	40	60	20	5	40	60	40	5	5	40	60	5	5	40	60	5
15	40	5	20	5	60	20	5	40	5	20	60	40	20	20	40	40	5	5	20	40
16	5	20	60	20	40	60	40	5	40	5	20	40	40	40	40	40	40	5	20	40
17	5	5	40	5	20	20	5	60	20	5	60	5	5	5	60	20	5	5	5	40
18	20	5	20	20	40	40	60	60	5	5	60	5	20	60	40	5	60	60	20	5
19	40	40	20	20	20	40	60	20	40	60	5	20	40	20	20	40	40	5	40	20
20	5	5	40	5	20	40	20	5	20	40	60	40	40	5	5	60	40	40	20	20

FIGURE 35 – Matrice d'échantillons à allocation optimale

— Les estimations

	ybar.opt	var.estim.ybaropt	ICinf	ICsup	IClong
1	23.875	18.941667	15.34468	32.40532	17.06063
2	28.125	22.625000	18.80211	37.44789	18.64577
3	27.750	12.377778	20.85432	34.64568	13.79137
4	17.875	9.880556	11.71406	24.03594	12.32187
5	32.250	16.600000	24.26435	40.23565	15.97129
6	31.125	15.230556	23.47584	38.77416	15.29833
7	29.750	21.977778	20.56143	38.93857	18.37714
8	27.750	20.866667	18.79671	36.70329	17.90658
9	28.500	16.544444	20.52773	36.47227	15.94455
10	22.000	17.766667	13.73850	30.26150	16.52301
11	36.375	23.175000	26.93948	45.81052	18.87104
12	29.125	16.419444	21.18290	37.06710	15.88420
13	26.125	12.480556	19.20075	33.04925	13.84851
14	22.000	14.488889	14.53940	29.46060	14.92120
15	34.000	17.344444	25.83725	42.16275	16.32549
16	25.500	16.400000	17.56260	33.43740	15.87479
17	27.250	19.866667	18.51388	35.98612	17.47224
18	18.750	14.311111	11.33531	26.16469	14.82937
19	29.625	17.475000	21.43159	37.81841	16.38682
20	28.875	15.886111	21.06295	36.68705	15.62409

FIGURE 36 – Matrice des estimations à allocation optimale

- La moyenne = 27.33125
- Les intervalles de confiance

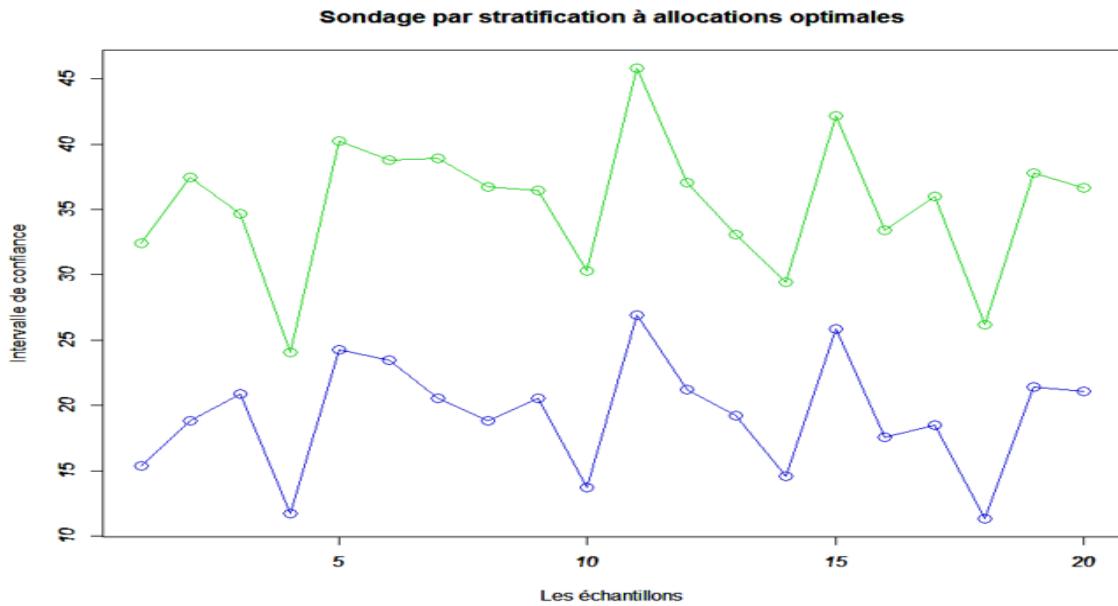


FIGURE 37 – Graphe des intervalles de confiance à allocations optimales

### 3.4 Sondage par grappes et à deux degrés

#### 3.4.1 Sondage par grappe

Cette méthode est utilisée généralement lorsqu'on ne dispose pas d'une base de sondage exhaustive. Elle consiste à : - Partitionner la population en  $M$  sous-population (unités primaires), - Sélectionner un échantillon d'unités primaires (les grappes), - Observer tous individus dans les unités sélectionnées.

Nous avons choisis de partitionner notre population selon la variable "CSP" ( catégorie socio professionnelle ) en 4 unités primaires.

— Les échantillons

	Ech1	Ech2	Ech3	Ech4	Ech5	Ech6	Ech7	Ech8	Ech9	Ech10
23	5	5	5	5	5	5	20	5	5	5
24	40	5	40	60	60	40	5	5	60	40
25	60	40	60	20	20	60	5	60	20	60
26	5	60	5	0	0	5	60	20	0	5
27	40	20	60	0	0	40	5	40	0	40
28	60	20	60	0	0	60	60	5	0	60
29	60	40	5	0	0	60	5	60	0	60
30	20	60	5	0	0	20	40	5	0	20
31	40	20	60	0	0	40	60	5	0	40
32	20	0	20	0	0	20	40	5	0	20
33	5	0	40	0	0	5	20	5	0	5
34	40	0	5	0	0	40	40	60	0	40
35	20	0	60	0	0	20	5	5	0	20
36	20	0	5	0	0	20	5	5	0	20
37	20	0	5	0	0	20	40	40	0	20
38	5	0	5	0	0	5	20	5	0	5
39	5	0	5	0	0	5	5	60	0	5
40	40	0	60	0	0	40	40	20	0	40
41	40	0	5	0	0	40	20	40	0	40
42	5	0	5	0	0	5	5	20	0	5
43	5	0	40	0	0	5	20	5	0	5
44	20	0	5	0	0	20	60	40	0	20

FIGURE 38 – Matrice d'échantillons issus de la méthode par grappe

— Les estimations

	YbarGrap	VarestGrap	BInf_IC	BSup_IC	LgIC
Ech1	27.6	205.21512	-0.4776497	55.67765	56.15530
Ech2	17.4	62.09513	1.9551093	32.84489	30.88978
Ech3	39.0	219.77853	9.9431388	68.05686	58.11372
Ech4	14.4	35.60640	2.7044647	26.09554	23.39107
Ech5	14.4	35.60640	2.7044647	26.09554	23.39107
Ech6	27.6	205.21512	-0.4776497	55.67765	56.15530
Ech7	25.8	77.31139	8.5663282	43.03367	34.46734
Ech8	36.0	198.62174	8.3770878	63.62291	55.24582
Ech9	14.4	35.60640	2.7044647	26.09554	23.39107
Ech10	27.6	205.21512	-0.4776497	55.67765	56.15530

FIGURE 39 – Matrice des estimations par la méthode des grappes

— La moyenne = 24.42

- 
- Les intervalles de confiance
- 

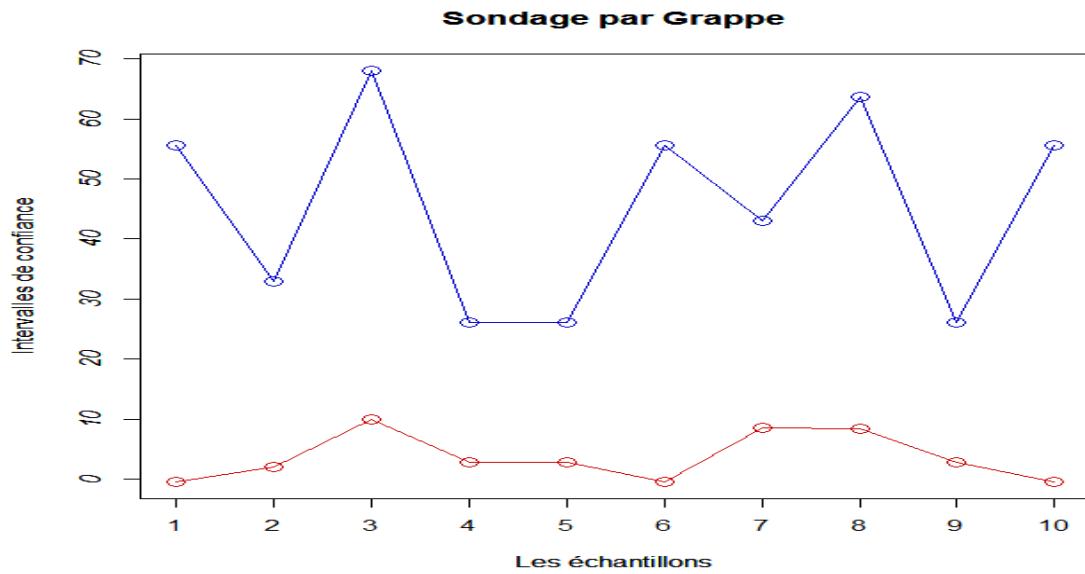


FIGURE 40 – Graphe des intervalles de confiance par la méthode des grappes

### 3.4.2 Sondage à deux degrés

La mise en œuvre de cette méthode est similaire à la précédente, seulement au lieu d'observer tous les individus dans les différentes grappes, on tire dans chaque grappe un échantillon de taille  $nh$ .

- 
- Les estimations
-

	ybar_chap	var_estim_ybar_Groupe	ICmin	ICmax	longueur_IC
1	26.47	1607.602	-52.11604	105.05604	157.1721
2	39.49	1607.602	-39.09604	118.07604	157.1721
3	24.09	1607.602	-54.49604	102.67604	157.1721
4	43.01	1607.602	-35.57604	121.59604	157.1721
5	23.72	1607.602	-54.86604	102.30604	157.1721
6	37.37	1607.602	-41.21604	115.95604	157.1721
7	24.00	1607.602	-54.58604	102.58604	157.1721
8	39.44	1607.602	-39.14604	118.02604	157.1721
9	26.18	1607.602	-52.40604	104.76604	157.1721
10	38.34	1607.602	-40.24604	116.92604	157.1721
11	28.01	1607.602	-50.57604	106.59604	157.1721
12	31.50	1607.602	-47.08604	110.08604	157.1721
13	48.03	1607.602	-30.55604	126.61604	157.1721
14	13.00	1607.602	-65.58604	91.58604	157.1721
15	16.36	1607.602	-62.22604	94.94604	157.1721
16	25.94	1607.602	-52.64604	104.52604	157.1721
17	26.75	1607.602	-51.83604	105.33604	157.1721
18	31.32	1607.602	-47.26604	109.90604	157.1721
19	40.54	1607.602	-38.04604	119.12604	157.1721
20	15.00	1607.602	-63.58604	93.58604	157.1721

FIGURE 41 – Matrice des estimations par le sondage à deux degrés

- La moyenne = 29.928
- Les intervalles de confiance

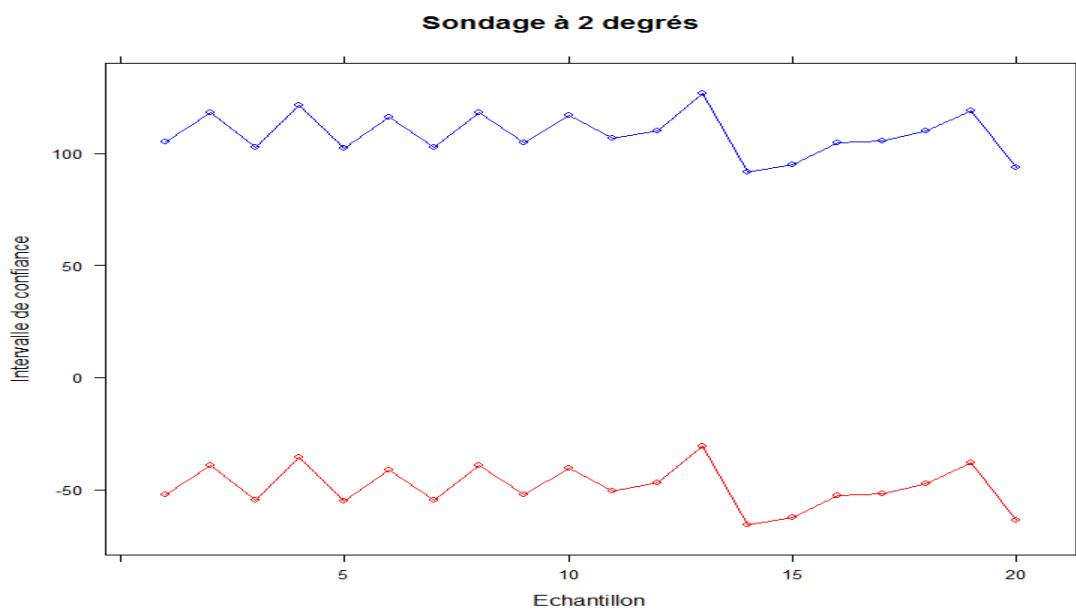


FIGURE 42 – Graphe des intervalles de confiance par le plan à deux degrés

## 4 Synthèse

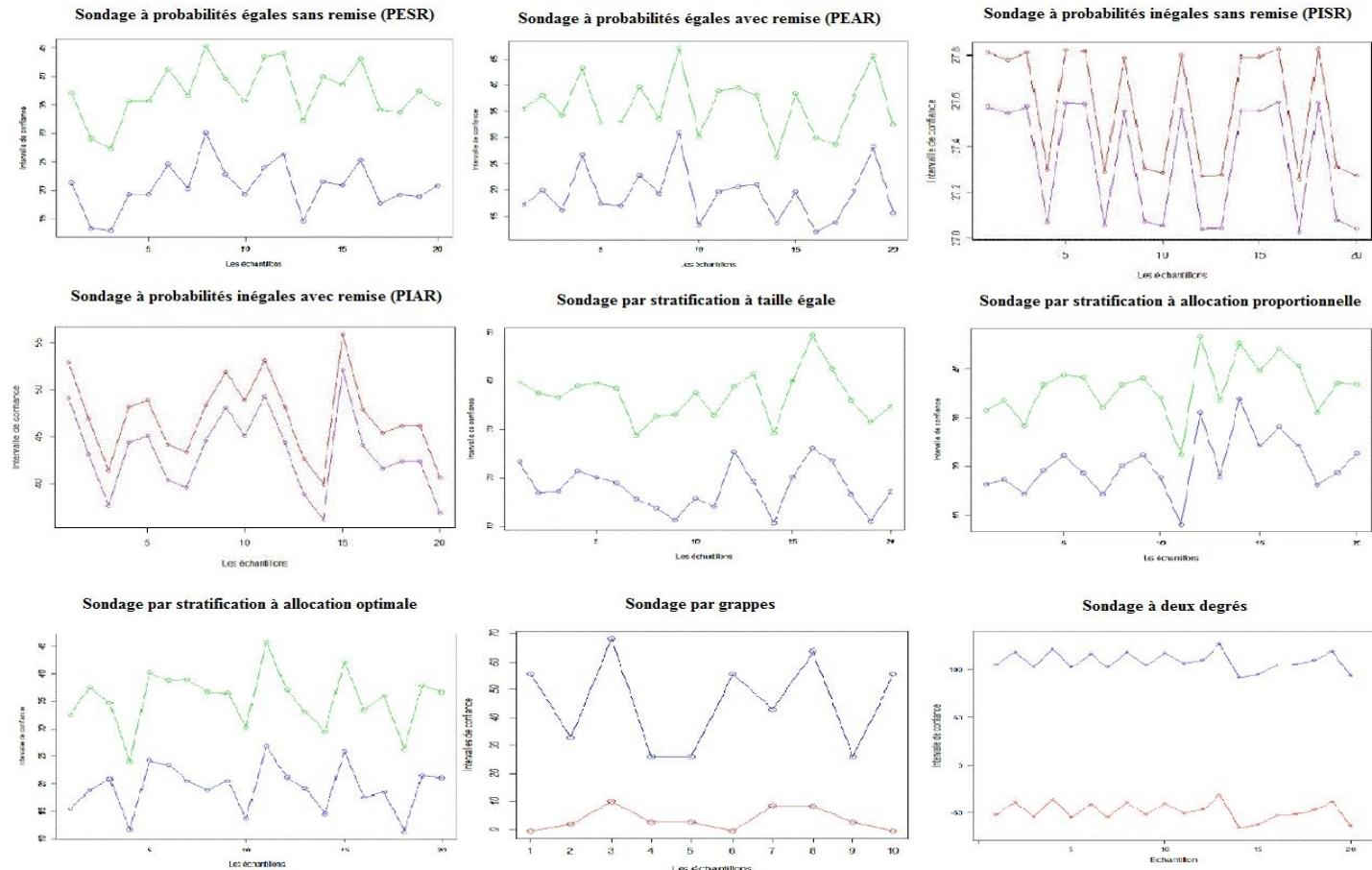


FIGURE 43 – Les différents Intervalles de confiance

	PESR	PEAR	PISR	PIAR	STRAT	PROP	OPT	GRAP	2DEG
MoygIC	16.634	16.956	3.786	0.233	19.048	15.918	16.004	41.735	57.172
MoyVarEst	28.908	27.381	27.381	27.381	27.381	28.400	27.331	24.420	20.928

FIGURE 44 – Matrice des moyennes et des IC pour chaque méthode de sondage

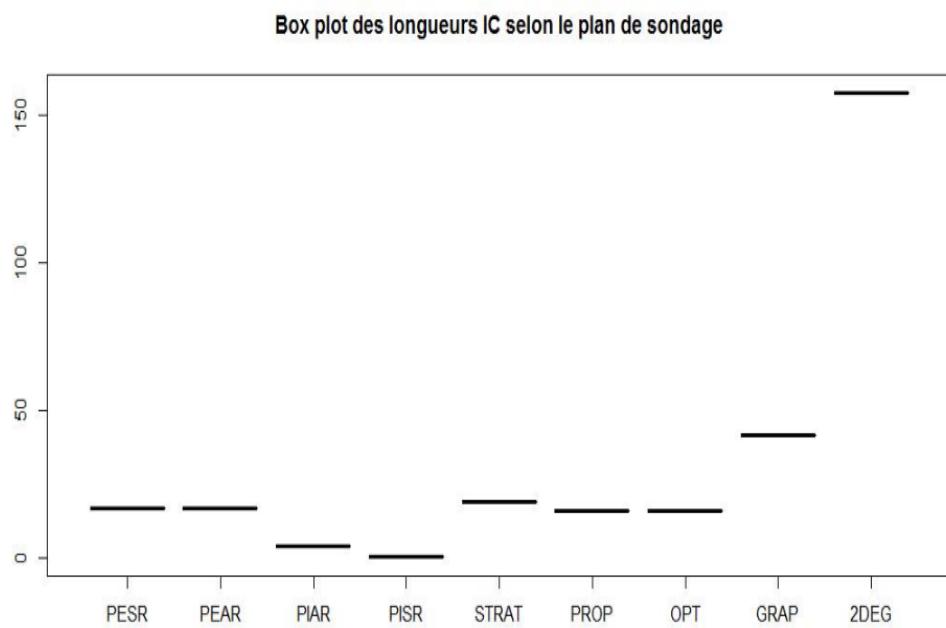


FIGURE 45 – Boîte à moustaches des longueurs IC0.95 par plan de sondage

---

## 5 Comparaison et commentaires

Toutes les méthodes ont fourni en moyenne des estimations de la moyenne non loin de la vraie valeur ( $=26.7$  DT) sauf celle à deux degré où on a obtenu une moyenne de 29.928 DT. Parmi les méthodes de sondage pratiquées, seul le sondage à probabilités inégales avec remise (PIAR) qui nous a permis d'obtenir des meilleurs résultats en termes de précision avec des longueurs IC en moyenne de 0.233.

Les sondages à probabilités égales ont été moins précis que ceux à probabilités inégales. Ce résultat était prévisible vu la grande différence entre les moyennes des intervalles de confiance de chaque méthode. (PISR de longueur 3.876 contrairement au PESR qui est de longueur 16.634).

D'une part les longueurs des intervalles de confiance issus des sondages à deux degrés sont relativement très élevées dont la moyenne converge vers 157.172 ce qui nous laisse conclure que cette méthode est la moins précise dans notre étude. Ce manque de précision est la contrepartie des erreurs d'échantillonnage liées à cette méthode.

D'une part la variance issue du sondage par grappe est très grande par rapport aux autres variances associées aux autres sondages donc le sondage par grappes n'est pas le plus judicieux pour notre enquête.

---

## 6 Redressement

Le redressement a pour but d'améliorer les estimations par utilisation de l'information auxiliaire. On part d'un estimateur simple (issu d'un sondage aléatoire simple) et on modifie les pondérations de l'échantillon pour obtenir l'estimation désirée. Parmi les 3 grandes méthodes de redressement, nous allons utiliser la post-stratification et l'estimation par le quotient.

### 6.1 Redressement par post stratification

La variable auxiliaire (VAX) utilisée est le Genre (à noter que la VAX doit être qualitative dans méthode), les effectifs par sexe constituent pour nous l'information auxiliaire. On tire un échantillon par un SAS, ensuite on le partitionne en m post strates selon notre variable auxiliaire. L'étape finale consiste à pondérer les données par l'information auxiliaires. Nous avons obtenu :

	<b>Post_St_F</b>	<b>Post_St_H</b>	<b>Ybar_redressé</b>
<b>1</b>	<b>29.66667</b>	<b>19.00000</b>	<b>27.00000</b>
<b>2</b>	<b>34.68750</b>	<b>32.50000</b>	<b>34.14062</b>
<b>3</b>	<b>34.09091</b>	<b>15.00000</b>	<b>29.31818</b>
<b>4</b>	<b>39.09091</b>	<b>23.33333</b>	<b>35.15152</b>
<b>5</b>	<b>27.00000</b>	<b>19.00000</b>	<b>25.00000</b>
<b>6</b>	<b>29.06250</b>	<b>8.75000</b>	<b>23.98438</b>
<b>7</b>	<b>25.38462</b>	<b>17.14286</b>	<b>23.32418</b>
<b>8</b>	<b>29.66667</b>	<b>11.00000</b>	<b>25.00000</b>
<b>9</b>	<b>26.56250</b>	<b>5.00000</b>	<b>21.17188</b>
<b>10</b>	<b>30.00000</b>	<b>22.50000</b>	<b>28.12500</b>

FIGURE 46 – Les estimations par post stratification

- Moyenne obtenue par post stratification :27.22157

## 6.2 Redressement par le quotient

Cette méthode nécessite une variable auxiliaire quantitative. Nous avons choisi la variable « age moyen »

	<b>moyxpi</b>	<b>moyVI</b>	<b>Yestim</b>
<b>Ech1</b>	<b>22.1</b>	<b>17.0</b>	<b>18.32308</b>
<b>Ech2</b>	<b>23.2</b>	<b>32.0</b>	<b>32.85517</b>
<b>Ech3</b>	<b>22.1</b>	<b>39.0</b>	<b>42.03529</b>
<b>Ech4</b>	<b>23.2</b>	<b>33.0</b>	<b>33.88190</b>
<b>Ech5</b>	<b>22.1</b>	<b>33.5</b>	<b>36.10724</b>
<b>Ech6</b>	<b>22.1</b>	<b>26.5</b>	<b>28.56244</b>
<b>Ech7</b>	<b>23.2</b>	<b>24.0</b>	<b>24.64138</b>
<b>Ech8</b>	<b>23.2</b>	<b>23.0</b>	<b>23.61466</b>
<b>Ech9</b>	<b>23.2</b>	<b>24.0</b>	<b>24.64138</b>
<b>Ech10</b>	<b>23.2</b>	<b>15.5</b>	<b>15.91422</b>

FIGURE 47 – Les estimations par redressement par le quotient

- Moyenne obtenue par redressement par le quotient : 28.05768

---

## 7 Conclusion

En conclusion, nous pouvons affirmer que cette enquête nous a permis d'estimer le budget moyen consacré pour la friperie pendant les 3 derniers mois à travers les différentes méthodes de sondage . Chacune de ces dernières nous a fourni un tas d'information avec un degré de précision que nous avons pu toucher à travers la longueur de l'intervalle de confiance. Nous pouvons constater que lorsque le nombre d'échantillon augmente, la moyenne de tous les estimateurs de chaque échantillon (20 dans notre cas) devient plus précise.

Nous pouvons constater aussi que le budget moyen varie entre 24.420 et 29.928 et c'était plus clair dans le sondage aléatoire simple sans remise qu'on peut le considérer comme le plus précis. Également pour le sondage aléatoire à probabilités inégales où nous avons obtenu presque le même résultat :la moyenne des estimateurs converge vers 27.32 DT, par contre, le sondage par grappe nous fait perdre la précision telle que les estimateurs de la moyenne s'éloignent de la vraie valeur de la moyenne ainsi que les intervalles de confiance deviennent plus larges.

Finalement, nous avons remarqué que la variable auxiliaire joue un rôle très important dans la précision des estimateurs, dans notre cas le genre , en faisant intercaler cette variable dans la stratification ,il est bien clair que la qualité de l'estimateur s'améliore par comparaison avec le sondage aléatoire simple(SAS).

---

## 8 Annexe

Sondage à probabilités égales  
PESR  
Tirage sans remise  
Initialisation des variables  
ech.pesr=matrix(0,nrow=20,ncol=20) matrice contenant les 20 échantillons  
ybar.pesr=NULL vecteur contenant les 20 estimateurs de la moyenne  
var.pesr=NULL vecteur contenant les variances pour les 20 échantillons  
varpesr.estim=NULL vecteur contenant les 20 estimateurs de la variance  
ICinf=NULL vecteur contenant les 20 borne inf des IC  
ICsup=NULL vecteur contenant les 20 borne sup des IC  
long.IC=NULL vecteur contenant la longueur des intervalles de confiance pour les 20 échantillons  
Tirage de 20 échantillons  
ech.pesr=replicate(20,sample(X\$Budget<sub>moy</sub>, 20, replace = FALSE, prob = NULL))  
nomcol=c("E1","E2","E3","E4","E5","E6","E7","E8","E9","E10","E11","E12","E13","E14","E15","E16","E17","E18","E19","E20")  
colnames(ech.pesr)=nomcol  
echant.pesr=data.frame(ech.pesr)  
View(echant.pesr)  
Résultats des 20 échantillons  
for (i in 1 :20)  
ybar.pesr[i]=mean(echant.pesr[,i])  
var.pesr[i]=var(echant.pesr[,i])  
varpesr.estim[i]=(1-f)\*var(echant.pesr[,i])/n  
ICinf[i]= ybar.pesr[i]-1.96\*sqrt(varpesr.estim[i])  
ICsup[i]= ybar.pesr[i]+1.96\*sqrt(varpesr.estim[i])  
long.IC[i]=ICsup[i]-ICinf[i]  
Représentation des résultats  
res.pesr=data.frame(ybar.pesr,var.pesr,varpesr.estim,ICinf,ICsup,long.IC)  
View(res.pesr)  
Représentation des intervalles de confiance  
plot(c(1 :20),ICinf,type="o", ylim= c( min(min(ICinf),min(ICsup)), max( max(ICinf),max(ICsup) ) ), main="SAS  
sans remise",xlab="Les échantillons",ylab="Intervalle de confiance",col="blue3",cex=1.5)  
par(new=TRUE)  
plot(c(1 :20),ICsup,type="o", ylim= c( min(min(ICinf),min(ICsup)), max( max(ICinf),max(ICsup) ) ),main="SAS  
sans remise",xlab="Les échantillons",ylab="Intervalle de confiance",cex=1.5,font=3,col="green3")  
La moyenne des 20 estimateurs  
moy.ybarpesr=mean(ybar.pesr)  
moy.ybarpesr  
Pear

---

---

Tirage avec remise

```

ech.pear=matrix(0,nrow=20,ncol=20) matrice contenant les 20 échantillons
ybar.pear=NULL vecteur contenant les 20 estimateurs de la moyenne
var.pear=NULL vecteur contenant les variances pour les 20 échantillons
varpear.estim=NULL vecteur contenant les 20 estimateurs de la variance
ICinf=NULL vecteur contenant les 20 borne inf des IC
ICsup=NULL vecteur contenant les 20 borne sup des IC
long.IC=NULL vecteur contenant la longueur des intervalles de confiance pour les 20
échantillons

```

Tirage de 20 échantillons

```

ech.pear=replicate(20,sample(X$Budgetmoy, 20, replace = TRUE, prob = NULL))
nomcol=
c("E1","E2","E3","E4","E5","E6","E7","E8","E9","E10","E11","E12","E13","E14","E15","E16","E17",
"")
colnames(ech.pear)=nomcol
echant.pear=data.frame(ech.pear)
View(echant.pear)

```

Résultats des 20 échantillons

```

for (i in 1 :20)
ybar.pear[i]=mean(echant.pear[,i])
var.pear[i]=var(echant.pear[,i])
varpear.estim[i]=(1-f)*var(echant.pear[,i])/n
ICinf[i]= ybar.pear[i]-1.96*sqrt(varpear.estim[i])
ICsup[i]= ybar.pear[i]+1.96*sqrt(varpear.estim[i])
long.IC[i]=ICsup[i]-ICinf[i]

```

Représentation des résultats

```

res.pear=data.frame(ybar.pear,var.pear,varpear.estim,ICinf,ICsup,long.IC)
View(res.pear)

```

Représentation des intervalles de confiance

```

plot(c(1 :20),ICinf,type="o", ylim= c( min(min(ICinf),min(ICsup)), max( max(ICinf),max(ICsup)
)), main="SAS
avec remise",xlab="Les échantillons",ylab="Intervalle de confiance",col="blue3",cex=1.5)
par(new=TRUE)
plot(c(1 :20),ICsup,type="o", ylim= c( min(min(ICinf),min(ICsup)), max( max(ICinf),max(ICsup)
)),main="SAS
avec remise",xlab="Les échantillons",ylab="Intervalle de confiance",cex=1.5,font=3,col="green3")

```

La moyenne des 20 estimateurs

```

moy.ybarpear=mean(ybar.pear)
moy.ybarpear

```

---

Sondage à probabilités inégales

-PIAR

library(lpSolve)

---

```

library(MASS)
library(sampling)
pk=inclusionprobabilities(y,n)/n
sum(pk)
v=c(0,cumsum(pk))
Echpiar=matrix(0,nrow=20,ncol=20) matrice qui va contenir les 20 échantillon
ICpiar=matrix(0,nrow=20,ncol=2) matrice qui va contenir les 20 IC
y.bar.piar=NULL vecteur qui va contenir les 20 estimateurs de la moyenne
var.piar=NULL vecteur qui va contenir les 20 estimateurs de la variance
Tirage de 20 échantillons
for (i in 1 :20)
s=NULL
for(j in 1 :20)
u=runif(1,0,1)
for(k in 2 :101)
if((v[k-1]==u)(v[k] > u))
s=c(s,k-1)
Echpiar[,i]=y[s]
y.bar.piar[i]=mean(Echpiar[,i])
var.piar[i]=var(Echpiar[,i])
Intervalles de confiance
for(i in 1 :20)
ICpiar[i,]=c(y.bar.piar[i]-1.96*sqrt(var(y.bar.piar)/20),y.bar.piar[i]+1.96*sqrt(var(y.bar.piar)/20))
longueur.piar=ICpiar[,2]-ICpiar[,1]
moy_y.bar.piar = mean(y.bar.piar)
Représentation des IC
plot(seq(1,20,by=1),ICpiar[,1],col="purple",ylim=c(min(min(ICpiar[,2]),min(ICpiar[,1])),max(max(IC
2]),max(ICpiar[,1]))),xlab= "Les échantillons",ylab="Intervalle de confiance")
lines(seq(1,20,by=1),ICpiar[,1],col="purple")
points(seq(1,20,by=1),ICpiar[,2],col="red")
lines(seq(1,20,by=1),ICpiar[,2],col="red")
Résultat
Les LA RÉVISION dans les 20 échantillons tirés
Echpiar
Les estimateurs de la moyenne dans les 20 échantillons tirés
y.bar.piar
Les estimateurs de la variance dans les 20 échantillons tirés
var.piar
Les 20 intervalles de confiances de la moyenne sont
ICpiar
Longueurs de Intervalles de Confiance
longueur.piar

```

---

---

moyenne de longueurs des intervalles de confiance  
 mean(longueur.piar)  
 Résumé  
 nomcolj-c("Ech1","Ech2","Ech3","Ech4","Ech5","Ech6","Ech7","Ech8","Ech9","Ech10","Ech11","Ech12")  
 colnames(Echpiar)j-nomcol  
 echantpiarj-data.frame(Echpiar)  
 Recappiarj-data.frame(y.bar.piar,var.piar,ICpiar[,1],ICpiar[,2],longueur)  
 colnames(Recappiar)j-c("Ybarpiar","VarYbarpiarEst","BInfIC","BSupIC","Longueur.piar")  
 row.names(Recappiar)j-nomcol  
 View(echantpiar)  
 View(Recappiar)  
 -PISR  
 N=100  
 n=20  
 library(lpSolve)  
 library(MASS)  
 library(sampling)  
 les probabilités d'inclusion  
 pi=inclusionprobabilities(y,n)  
 sum(pi)  
 Initialisation des variables  
 ybarPISR=NULL vecteur qui va contenir les 20 estimateurs de la moyenne  
 varPISR=NULL vecteur qui va contenir les 20 estimateurs de la variance  
 IC=matrix(0,nrow =20,ncol=2)matrice qui va contenir les 20 IC(Intervalle de Confiance)  
 echp ISR=matrix(0,20,20) matrice qui va contenir les 20 échantillons  
 Tirage de 20 échantillons  
 for(i in 1 :20)  
 pi=inclusionprobabilities(y,20)  
 sp=UPsystematic(pi)  
 ech=getdata(y,sp)  
 echp ISR[,i]=ech[,2]  
 pk=getdata(pi,sp)  
 tybarPISR=sum(ech[,2]/pk)  
 ybarPISR[i]=1/100\*tybarPISR  
 varPISR[i]=1/10000\*sum((1-pk[,2])\*(ech[,2]/pk[,2])^2)  
 Intervalle de Confiance  
 for(i in 1 :20)  
 IC[i,]=c(ybarPISR[i]-1.96\*sqrt(var(ybarPISR)/20),ybarPISR[i]+1.96\*sqrt(var(ybarPISR)/20))  
 longueur.pisr=IC[,2]-IC[,1]  
 moy\_y.bar.pisr = mean(ybarPISR)  
 Résultat  
 les 20 échantillons tirés

---

---

echpisr  
 Les estimateurs de la moyenne dans les 20 échantillons tirés  
 ybarPISR  
 Les estimateurs de la variance dans les 20 échantillons tirés  
 varPISR  
 Les 20 intervalles de confiances de la moyenne sont  
 IC  
 Longueurs des Intervalles de Confiance  
 longueur.pisr  
 moyenne de longueurs des Intervalles de confiance  
 mean(longueur.pisr)  
 Représentation des IC  
 plot(seq(1,20,by=1),IC[,1],col="purple",ylim=c(min(min(IC[,2]),min(IC[,1])),max(max(IC[,2]),max(IC[,1]))),xlab= "Les échantillons",ylab="Intervalle de confiance")  
 lines(seq(1,20,by=1),IC[,1],col="purple")  
 points(seq(1,20,by=1),IC[,2],col="red")  
 lines(seq(1,20,by=1),IC[,2],col="red")  
 La moyenne des 20 estimateurs de la moyenne  
 moy\_y.bar.pisr  
 Résumé  
 nomcol=j-c("Ech1","Ech2","Ech3","Ech4","Ech5","Ech6","Ech7","Ech8","Ech9","Ech10","Ech11","Ech12")  
 colnames(echpisr)=nomcol  
 echantpisr=data.frame(echpisr)  
 recappisr=data.frame(ybarPISR,varPISR,IC[,1],IC[,2],longueur.pisr)  
 colnames(recappisr)=c("Ybarpisr","VarYbarpisr<sub>Est</sub>","BInf<sub>I</sub>C","BSup<sub>I</sub>C","Longueur<sub>pisr</sub>")  
 row.names(recappisr)=nomcol  
 View(echantpisr)  
 View(recappisr)

---

Stratification  
 -Strates à tailles égales  
 View(X)  
 attach(X)  
 N=100  
 n=20  
 library(MASS)  
 library(lpSolve)  
 library(sampling)  
 Strates de même taille  
 Initialisation des variables  
 echant.strate=matrix(0,nrow=20,ncol=20)  
 ybar.strate=NULL

---

---

```

ybar.pesr.h= NULL
ybar.pesr.f=NULL
ybar.strate=NULL
N=100 nombre d'observation dans la base
Nf=75 taille de la strate des femmes
Nh=25 taille de la strate des hommes
nh=10 nombre d'échantillon sur chaque strate
ICinf=NULL contient les bornes inférieurs des intervalles de confiance
ICsup=NULL contient les bornes supérieures des intervalles de confiance
IClong=NULL contient la longueur de chaque intervalle de confiance pour les 20 échantillons
Préparation de l'échantillon
y=X$Budgetmoy
for(i in 1 :20)
strate=strata(X,stratanames="Sexe",size=c(nh,nh),method="srswor")
echant.strate[,i]=y[strate$IDuit]
nomcol=c("Ech1","Ech2","Ech3","Ech4","Ech5","Ech6","Ech7","Ech8","Ech9","Ech10","Ech11","Ech12","Ech13","Ech14","Ech15","Ech16","Ech17","Ech18","Ech19","Ech20")
colnames(echant.strate)=nomcol
View(echant.strate)
Calcul des résultats sur les échantillons
for (i in 1 :20)
ybar.pesr.h=mean(echant.strate[1 :10,i])
ybar.pesr.f=mean(echant.strate[11 :20,i])
ybar.strate[i]=(Nh/N)*ybar.pesr.h + (Nf/N)*ybar.pesr.f
s2c.h=var(echant.strate[1 :10,i])
s2c.f=var(echant.strate[11 :20,i])
var.estim.strate[i]= ( ((Nh2)/nh) * (1 - nh/Nh) * s2c.h + ((Nf2)/nh) * (1 - nh/Nf) * s2c.f)/(N2)
ICinf[i]= ybar.strate[i]-1.96*sqrt(var.estim.strate[i])
ICsup[i]= ybar.strate[i]+1.96*sqrt(var.estim.strate[i])
IClong[i]=ICsup[i]-ICinf[i]
res.strate=data.frame(ybar.strate,var.estim.strate,ICinf,ICsup,IClong)
View(res.strate)
La moyenne des 20 estimateurs
moy.ybarstrate=mean(ybar.strate)
moy.ybarstrate
Représentation des intervalles de confiance
plot(c(1 :20),ICinf,type="o", ylim= c( min(min(ICinf),min(ICsup)), max( max(ICinf),max(ICsup))), main="Sondage par stratification à taille égale",xlab="Les échantillons",ylab="Intervalle de confiance",col="blue3",cex=1.5)

```

---

---

```

par(new=TRUE) plot(c(1 :20),ICsup,type="o", ylim= c( min(min(ICinf),min(ICsup)),
max( max(ICinf),max(ICsup)
 )),main="Sondage par stratification à taille égale",xlab="Les échantillons",ylab="Intervalle
de
confiance",cex=1.5,font=3,col="green3")
- Strates à allocations proportionnelles
View(X)
attach(X)
N=100
n=20
library(MASS)
library(lpSolve)
library(sampling)
Nh=25 taille de la strate des hommes
Nf=75 taille de la strate des femmes
nh=round((n/N)*Nh) taille de l'échantillon des hommes
nf=round((n/N)*Nf) taille de l'échantillon des femmes
f= n/N
Initialisation des variables
echant=matrix(0,nrow=20,ncol=20) matrice contenant les 20 échantillons
ybar.prop=NULL contient la moyenne de chaque échantillon
var.estim.ybarprop=NULL contient la variance estimée de chaque échantillon
ICinf=NULL contient les bornes inférieurs des intervalles de confiance
ICsup=NULL contient les bornes supérieures des intervalles de confiance
IClong=NULL contient la longueur de chaque intervalle de confiance pour les 20 échantillons
Préparation de l'échantillon
y= X$Budgetmoy
data.h= y[X$Sexe=="Homme"]
data.f= y[X$Sexe=="Femme"]
for(i in 1 :20)
echant.h=sample(data.h,nh,replace=TRUE)
echant.f=sample(data.f,nf,replace=TRUE)
echant[,i]=c(echant.h,echant.f)
nomcol=c("Ech1","Ech2","Ech3","Ech4","Ech5","Ech6","Ech7","Ech8","Ech9","Ech10","Ech11","Ech12")
colnames(echant)=nomcol
View(echant)
Calcul des résultats sur les échantillons
for(i in 1 :20)
ybar.prop[i]=mean(echant[,i])
s2.h=var(echant[1 :10,i])
s2.f=var(echant[11 :20,i])
var.estim.ybarprop[i]=((Nh2)/nh) * (1 - f) * s2.h + ((Nf2)/nf) * (1 - f) * s2.f)/(N2)

```

---

---

```

ICinf[i]= ybar.prop[i]-1.96*sqrt(var.estim.ybarprop[i])
ICsup[i]= ybar.prop[i]+1.96*sqrt(var.estim.ybarprop[i])
IClong[i]=ICsup[i]-ICinf[i]
res.prop=data.frame(ybar.prop,var.estim.ybarprop,ICinf,ICsup,IClong)
View(res.prop)
La moyenne des 20 estimateurs
moy.ybarprop=mean(ybar.prop)
moy.ybarprop
Représentation des intervalles de confiance
plot(c(1 :20),ICinf,type="o", ylim= c( min(min(ICinf),min(ICsup)), max( max(ICinf),max(ICsup))
)), main="Sondage par stratification à allocations proportionnelles",xlab="Les échantillons",ylab="Interval
de confiance",col="blue3",cex=1.5)
par(new=TRUE)
plot(c(1 :20),ICsup,type="o", ylim= c( min(min(ICinf),min(ICsup)), max( max(ICinf),max(ICsup)
)),main="Sondage par stratification à allocations proportionnelles",xlab="Les échantillons",ylab="Interval
de confiance",cex=1.5,font=3,col="green3")
- Stratification à allocation optimale
View(X)
attach(X)
N=100
n=20
Nh=25
Nf=75
library(MASS)
library(lpSolve)
library(sampling)
y= X$Budgetmoy
data.h= y[X$Sexe=="Homme"] c'est la strate "homme"
data.f= y[X$Sexe=="Femme"] c'est la strate "femme"
Calcul de la taille d'échantillon sur chaque strate
sig2c.h=var(data.h) variance de la strate "Homme"
sig2c.f=var(data.f) variance de la strate "Femme"
nh=round(n*(Nh*sqrt(sig2c.h))/(Nh*sqrt(sig2c.h)+Nf*sqrt(sig2c.f))) taille de l'échantillon
des hommes
nf=round(n*(Nf*sqrt(sig2c.f))/(Nh*sqrt(sig2c.h)+Nf*sqrt(sig2c.f))) taille de l'échantillon
des femmes
fh=nh/Nh taux de sondage pour la strate "Homme"
ff=nf/Nf taux de sondage pour la strate "Femme"
Initialisation des variables
echant=matrix(0,nrow=20,ncol=20) matrice contenant les 20 échantillons
ybar.opt=NULL contient la moyenne de chaque échantillon

```

---

---

```

var.estim.ybaropt=NULL contient la variance estimée de chaque échantillon
ICinf=NULL contient les bornes inférieures des intervalles de confiance
ICsup=NULL contient les bornes supérieures des intervalles de confiance
IClong=NULL contient la longueur de chaque intervalle de confiance pour les 20 échantillons
Préparation de l'échantillon
for(i in 1 :20)
echant.h=sample(data.h,nh,replace=TRUE)
echant.f=sample(data.f,nf,replace=TRUE)
echant[,i]=c(echant.h,echant.f)
nomcol=c("Ech1","Ech2","Ech3","Ech4","Ech5","Ech6","Ech7","Ech8","Ech9","Ech10","Ech11","Ech12")
colnames(echant)=nomcol
View(echant)
Calcul des résultats sur les échantillons
for(i in 1 :20)
moy.h=mean(echant[1 :10,i])
moy.f=mean(echant[11 :20,i])
ybar.opt[i] = ( Nh*moy.h + Nf*moy.f )/N
var.h=var(echant[1 :10,i])
var.f=var(echant[11 :20,i])
var.estim.ybaropt[i]=( (Nh2)*((1-fh)/nh)*var.h+(Nf2)*((1-ff)/nf)*var.f)/(N2)
ICinf[i]= ybar.opt[i]-1.96*sqrt(var.estim.ybaropt[i])
ICsup[i]= ybar.opt[i]+1.96*sqrt(var.estim.ybaropt[i])
IClong[i]=ICsup[i]-ICinf[i]
res.opt=data.frame(ybar.opt,var.estim.ybaropt,ICinf,ICsup,IClong)
View(res.opt)
La moyenne des 20 estimateurs
moy.ybaropt=mean(ybar.opt)
moy.ybaropt
Représentation des intervalles de confiance
plot(c(1 :20),ICinf,type="o", ylim= c( min(min(ICinf),min(ICsup)), max( max(ICinf),max(ICsup) )), main="Sondage par stratification à allocations optimales",xlab="Les échantillons",ylab="Intervalle de confiance",col="blue3",cex=1.5)
par(new=TRUE)
plot(c(1 :20),ICsup,type="o", ylim= c( min(min(ICinf),min(ICsup)), max( max(ICinf),max(ICsup) )),main="Sondage par stratification à allocations optimales",xlab="Les échantillons",ylab="Intervalle de confiance",cex=1.5,font=3,col="green3")

```

---

Sondage par grappes et à deux degrés

Sondage par grappes

sd=as.data.frame(datafinal)

library(MASS)

library(lpSolve)

---

```

library(sampling)
N=100
n=20
NG=as.vector(table(sd$Agemoy))
M=length(NG)
m=2
yGrap = NULL
vGrap = NULL
vyGrap = NULL
lgICGrap = NULL
IC=matrix(nrow=10,ncol=2)
data=sd
res=j-matrix(0,nrow=2*max(table(sd$Agemoy)),ncol=10) matrice qui contenir les individus selectiones
for (i in 1 :10)
sg=cluster(data,clustername=c("Agemoy"),size=2,method="srswor")
pik=sg$Prob
ech=getdata(data,sg)
y1=ech$Budgetmoy
res[1 :length(y1),i]=y1
gr=unique(sg$Agemoy)
tg = NULL
ng=NULL
for(j in 1 :2)
tg[j] = sum(y1[ech$Agemoy == gr[j]])
ng[j]=length(sg$Agemoy[sg$Agemoy==gr[j]])
pi=unique(pik)
tGrap = sum(tg)/pi
yGrap[i] = tGrap/N
differencecarre = NULL
for(j in 1 :2) in 1 :m
differencecarre[j] = (tg[j] - (tGrap/ng[j]))2
vyGrap[i] = 2 * (2/3) * (1/10000) * sum(differencecarre)
((M/m)2) * ((M - m)/(M1)) * (1/N2) * sum(differencecarre)
IC[i,1]=yGrap[i] - 1.96 * sqrt(vyGrap[i])
IC[i,2]=yGrap[i] + 1.96 * sqrt(vyGrap[i])
lgIC=IC[,2]-IC[,1]
echgrap=res
lgIC=IC[,2]-IC[,1]
nomcol=j-c("Ech1","Ech2","Ech3","Ech4","Ech5","Ech6","Ech7","Ech8","Ech9","Ech10","Ech11","Ech19","Ech20")
colnames(echgrap)=j-c("Ech1","Ech2","Ech3","Ech4","Ech5","Ech6","Ech7","Ech8","Ech9","Ech10")

```

---

---

```

echgrap=data.frame(echgrap)
recapgrapj-data.frame(yGrap,vyGrap,IC[, 1], IC[, 2], lgIC)
colnames(recapgrapj)-c("YbarGrap","VarestGrap","BInfIC","BSupIC","LgIC")
row.names(recapgrapj)-nomcol
View(echgrap)
View(recapgrap)
mean(yGrap)
plot(c(1 :10),IC[,1],ylim=c(min(min(IC[,2]),min(IC[,1])),max(max(IC[,2]),max(IC[,1]))),type="o",xaxt
"n",
main="Sondage par Grappe",xlab="Les échantillons",ylab="Intervalles de confiance",col="red3",cex=
par(new=TRUE)
plot(c(1 :10),IC[,2],ylim=c(min(min(IC[,2]),min(IC[,1])),max(max(IC[,2]),max(IC[,1]))),type="o",xaxt
"n",main="Sondage par Grappe",xlab="Les échantillons",ylab="Intervalles de confiance",cex=1.5,font
axis(1,at=c(1 :10))
Sondage à deux degrés
library(MASS)
library(lpSolve)
library(sampling)
connexion=read.table(file.choose(),header=T,sep=';')
Ni=as.vector(table(connexion$CSPm))
ybarChap = NULL
varestimybarGrappe = NULL
icSgmin=NULL
icSgmax=NULL
M=4
m=2
for(l in
1 :20)Sg=cluster(connexion,clustername=c("CSPm"),size=2,method="srswor")
pik=Sg$Prob
EchSg = getdata(connexion,Sg)
n=dim(Sg)[1]
yech = EchSg$Budgetmoy
gr=unique(Sg$CSPm)
ty=c(0,0)
nh=round(Ni/5)
for (i in 1 :2)
p2=srswor(nh[gr[i]],Ni[gr[i]])
grappe=EchSg$Budgetmoy[EchSg$CSPm == gr[i]]
ech2=getdata(EchSg$Budgetmoy[EchSg$CSPm == gr[i]],p2)
yech2 = ech2[, 2]
if(nh[gr[i]] !=0) ty[i]=sum(yech2) * Ni[gr[i]]/nh[gr[i]] * (M/m)
ss=c(0,0)

```

---

---

```

for(q in 1 :Ni[gr[i]])
for(p in 1 :Ni[gr[i]])
if(p !=q)ss[i]=ss[i]+(grappe[q]*grappe[p]*(M2) * (Ni[gr[i]]2) * (11/(Ni[gr[i]] * M)))+
((grappe[q]/nh[gr[i]]/Ni[gr[i]])2 * (1 - 1/(M * Ni[gr[i]])))
ybarchap[l] = sum(ty)/100
Niech = as.vector(Ni[gr])
niech = round(Niech/5)
v1=sum(((ty*3)2) * 2/3) + 2 * ((ty[1]/m/M) * (ty[2]/m/M) * (1 - 1/(M * m)))
v2=sum(ss)
varestimybarGrappe = (v1 + v2)/10000
ICmin=-1.96*sqrt(varestimybarGrappe) + ybarchap
ICmax=+1.96*sqrt(varestimybarGrappe) + ybarchap
longueurIC = ICmax - ICmin
SG2=data.frame(ybarchap, varestimybarGrappe, ICmin, ICmax, longueurIC
)
s=1 :20
sg2=mean(SG2$ybarchap)
sg2
SG2
s=1 :20
h1=SG2$icSgmin
h2=SG2$icSgmax
library(lattice)
xyplot(h2+h1 ~ s,col=c("blue","red"),main="Sondage à 2 degrés ",ylab="Intervalle de
confiance",xlab="Echantillon",type='o')
View(SG2)

```

---

Synthèse

nombox=j-c("PESR","PEAR","PIAR","PISR","STRAT","PROP","OPT","GRAP","2DEG")

nombox

w=cbind(16.63391,16.9558755,3.786457,0.2330497,19.04836,15.91832,16.09391,41.73558,157.1721)

longueur de

chacun des intervalles de confiance par méthode

boxplot(w,names=nombox,main="Box plot des longueurs IC selon le plan de sondage",col=c("blue","red","gray","green","yellow","turquoise3","violetred4","magenta","cyan"))

MoylgIC=j-c(16.63391,16.9558755,3.786457,0.2330497,19.04836,15.91832,16.09391,41.73558,157.1721)

valeur de

la moyenne lgIC par méthode

Moyybar=c(28.9875,27.38125,27.38125,27.38125,27.38125,28.4,27.33125,24.42,29.928) Ybar

estimé moyenne

correspondant aux valeur de MoylgIC

recapprojetj-data.frame(MoylgIC,Moyybar)

---

```

recapprojet=t(recapprojet),3)
colnames(recapprojet)=c("PROP","OPT","STRAT","PESR","PEAR","PIAR","PISR","GRAP","2DE")
rownames(recapprojet)=c("MoylgIC","MoyYbarEst")
View(recapprojet)

Redressement
Redressement par post-stratification
data=as.data.frame(sd)
T=data[, "Budgetmoy"]
C=data[, "Sexem"] = Sexe
echt=vector("length"=20,"numeric")
ech=vector("length"=20,"numeric")
sech=vector("length"=20,"numeric")
yf=NULL
yh=NULL
ybarchap=NULL
p=1 :100
for(j in 1 :10)
ech=sample(p,20,replace=FALSE)
for (i in 1 :20)
k=ech[i]
echt[i]=T[k] echantillon
sech[i]=C[k] sexe des individus selectionner
echanf=echt[sech==2] la post strate des femmes
echanh=echt[sech==1] la post strate des hommes
nbf=length(echanf)
nbh=length(echanh)
NBf=sum(C==2) taille réelle de la strate de femmes
NBh=sum(C==1) taille réelle de la strate de hommes
yf[j]=mean(echanf)
yh[j]=mean(echanh)
ybarchap[j]=(NBf*yf[j]+NBh*yh[j])/100 la moyenne redrésséee
recapredr1=data.frame(yf,yh,ybarchap)
colnames(recapredr1)=c('PostStF','PostStH','Ybarredrésséé')
rownames(recapredr1)=c("Moyenne")
View(recapredr1)
mean(ybarchap)

Redressement par quotient
library(lpSolve) library(MASS) library(sampling)
moyxpi=NULL moyVI=NULL Yestim=NULL
for (j in 1 :10)
s=rswor(10,100) s echy=X$Budgetmoy[s == 1]echymoyVI[j] = mean(echy)moyennedelaVI dans l'éch

```

---

---

```
echx=X$Age_moy[s == 1]moyxpi[j] = mean(echx)moyennedelaVAXdansl'échantillonmoyxpi[j]
Xbar=mean(X$Age_moy)moyennedelaVAXdanslapopulationXbar = 23.82
Yestim[j]=(Xbar*moyVI[j])/moyxpi[j]
recapredj-data.frame(moyxpi,moyVI,Yestim) nomcol=c("Ech1","Ech2","Ech3","Ech4","Ech5","Ech6",
rownames(recapred)=nomcol View(recapred)
mean(Yestim)
```