



République Tunisienne
Ministère de l'Enseignement Supérieur et de la
Recherche Scientifique
Université de Carthage
Ecole Supérieure de la Statistique et de l'Analyse
de l'Information



PROJET DE SONDAGE

**Enquête sur le budget mensuel alloué à la pratique du sport
chez les tunisiens âgés de 15ans et plus dans les différentes
salles du sport suivantes : Star Fitness Gym - El Aouina,
Fitness Factory - Fouchena et Doctor Gym - La Soukra**

Réalisé par :

BACCOUCH Ismail

KHORCHANI Amal

SBAA Syrine

Superviseuse : Mme Ouaili Héra

Année Universitaire : 2022/2023

Remerciements

Ce travail ne serait pas fructueux et n'aurait pas été rendu possible sans l'aide et les conseils de Mme Ouaili Héla. Nous tenons très sincèrement à la remercier pour sa disponibilité, sa patience, ses explications et son encadrement enrichissant lors de la réalisation de ce projet. Notre gratitude aux enquêtés pour leur contrition dans la réalisation de ce travail. Nos remerciements vont également à nos aînés pour leur soutien durant la réalisation de ce projet. Finalement, un grand merci à tous ceux qui, de près ou de loin, nous soutiennent dans notre parcours académique

Table des matières

Remerciements	i
Table des figures	iv
Table des tables	v
Introduction	vi
1 Les généralités	1
1.1 Le Sondage	1
1.1.1 Définition	1
1.1.2 Les principes des sondages	1
1.1.3 Vocabulaire	1
1.1.4 Les étapes de mise en oeuvre d'un sondage	2
1.2 Les méthodes des sondages	2
2 Fiche d'enquête	6
2.1 Période d'enquête	6
2.2 Enquêteurs	6
2.3 Technique d'enquête	6
2.4 Traitement de données	6
2.5 Population source	6
2.6 Population cible	6
2.7 Variable d'intérêt et sa signification	7
2.8 Collecte des données	7
2.9 Questionnaire	7
2.10 Base des données	8
3 Resultats	9
3.1 Analyse des statistiques descriptives	9
3.1.1 Sexe	9
3.1.2 Age	10
3.1.3 Catégorie socio-professionnelle	10
3.1.4 Statut matrimonial	11
3.1.5 Activités sportives pratiquées	12

3.1.6	Heures	12
3.1.7	Endroit	13
3.1.8	Autres dépenses	13
3.1.9	Budget mensuel alloué à la pratique sport	14
3.2	Traitement des données	14
3.2.1	Sondage aléatoire simple :le SAS	14
3.2.2	Sondage aléatoire à probabilité inégales	19
3.2.3	La stratification	23
3.2.4	Sondage à plusieurs degrés	29
3.2.5	Synthèse	33
3.2.6	Redressement	33
Annexe		34

Table des figures

2.1	Base de données	8
3.1	Répartition des individus selon le sexe	9
3.2	Répartition des individus selon l'âge	10
3.3	Répartition des individus selon leur catégorie socio-professionnelle	10
3.4	Répartition des individus selon leur statut matrimonial	11
3.5	Répartition des individus selon leurs activités sportives	12
3.6	Répartition des individus selon les heures passées	12
3.7	Répartition des individus selon leurs endroits	13
3.8	Répartition des individus selon leurs dépenses supplémentaires	13
3.9	Répartition des individus selon leur budget mensuel	14
3.10	Echantillons PESR	15
3.11	Intervalle de confiance pour PESR	16
3.12	Echantillon PEAR	17
3.13	Intervalles de confiance pour PEAR	18
3.14	Echantillon PISR	19
3.15	Intervalle de confiance pour PISR	20
3.16	Echantillon PIAR	21
3.17	Intervalle de confiance pour PIAR	22
3.18	Echantillon pour strate de même taille	23
3.19	– Intervalle de confiance pour stratification avec strates de même taille	24
3.20	Echantillon pour stratification à allocations proportionnelles	25
3.21	Intervalle de confiance pour stratification à allocations proportionnelles	26
3.22	Echantillon pour stratification à allocations optimales	27
3.23	Intervalle de confiance pour stratification à allocation optimales	29
3.24	Echantillon pour sondage par grappe	30
3.25	Intervalle de confiance pour sondage par grappe	31
3.26	Echantillon pour sondage à deux degrés	31
3.27	Intervalle de confiance pour sondage à deux degrés	32

Table des tables

3.1 **Tableau récapitulatif** 33

Introduction

Aujourd'hui, un mode de vie sain est de plus en plus populaire parmi les jeunes et les personnes âgées. Les gens sont devenus plus conscients de leur santé. L'industrie de la gym et du fitness est florissante dans le pays. Dans une certaine mesure, le mérite en revient aux acteurs et aux influenceurs(ses), qui montrent à l'écran leurs corps et leurs muscles bien entretenus et inspirent les gens.

C'est pour cette raison que nous avons choisi de mener une enquête auprès d'un **échantillon de 100 tunisiens intéressés par le sport** et adoptant un **style de vie sain** afin de mesurer leur budget consacré au sport et ainsi analyser les principaux facteurs influençant cette grandeur.

Dans les lignes qui suivent, nous présenterons les résultats obtenus à partir du traitement des données recueillies selon les différentes méthodes vues lors du cours de sondage.

Chapitre 1

Les généralités

1.1 Le Sondage

1.1.1 Définition

En général, le sondage c'est l'action de sonder un milieu quelconque, une cavité, une eau, un sol c'est-à-dire une technique d'exploration d'un milieu. Statistiquement, on appelle sondage 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. Les sondages les plus connus du grand public portent sur des populations humaines. Ce sont en particulier les sondages d'opinion réalisés par des entreprises de sondage. Mais il ne s'agit là que d'une application particulière de la technique du sondage, dont l'usage est beaucoup plus général (voir aussi échantillonnage).

1.1.2 Les principes des sondages

Pour réaliser un sondage juste et adapté à la population visée, il faut choisir la technique d'échantillonnage adéquate. Il existe en effet de nombreuses méthodes permettant de créer un échantillon représentatif de la population mère. Il y a dans un premier temps les techniques probabilistes puis les techniques non probabilistes. Dans le premier cas, les individus de la population sont tirés au hasard et ont par conséquent tous une probabilité, égale ou non, de figurer dans l'échantillon. Dans le cas d'un échantillonnage non probabiliste, tous les individus n'ont pas forcément une probabilité non nulle de figurer dans l'échantillon.

1.1.3 Vocabulaire

- **La population cible** : c'est la population¹ pour laquelle l'information est requise. Elle est définie suivant quatre (4) caractéristiques : la nature, les caractéristiques intrinsèques, la localisation géographique et la date.

1. Liste exhaustive d'individus sans ambiguïté $P = c_1, \dots, c_N$

- **La population source** : est l'ensemble des individus à partir duquel sera constitué l'échantillon auprès duquel seront collectées les données.
- **La variable d'intérêt** : l'information à collecter est toujours reliée à un caractère qualitatif ou quantitatif : c'est la variable d'intérêt.
- **Echantillon, base de sondage, taux de sondage** : un sondage est un recueil des valeurs d'une partie (appelée n-échantillon car d'effectif n) de la population dite de base de sondage. On appelle **taux de sondage** le rapport $\frac{n}{N}$.

1.1.4 Les étapes de mise en oeuvre d'un sondage

1. Spécification des objectifs poursuivis et de la manière d'y parvenir : définition exacte des variables d'intérêt et de la population cible, rechercher les informations auxiliaires, spécifier le mode de collecte des données.
2. Construction d'une procédure permettant l'accès à toutes les unités de la population : labellisation de chaque individus garantissant ainsi une identification non ambiguë et un repérage des éventuelles doubles prises en compte.
3. Conception du questionnaire ou mise au point de l'interview
4. Construction du plan de sondage (méthode utilisée pour la sélection des unités au sein de la base de sondage). Le plan est dit probabiliste si la procédure de sélection est aléatoire ; sinon il est empirique.
5. Collecte des informations : l'enquête.
6. Dépouillement et contrôle de la qualité des données (codage, création d'un fichier de données et contrôle de la cohérence).
7. Estimation des paramètres : il s'agit d'estimer une fonction de la **variable d'intérêt** $\theta = f(Y_1, \dots, Y_N)$.
8. Evaluation des résultats : détection des risques d'erreurs à tous les niveaux de la procédure.
9. Publication des résultats (si cela l'importe).

1.2 Les méthodes des sondages

Quatre(4) méthodes seront utilisées dans notre travail

Le sondage aléatoire simple : SAS

Si tous les échantillons de même taille ont la même probabilité d'être sélectionnés alors le sondage est dit aléatoire simple. **Conséquence** : Tous les individus ont la même probabilités d'inclusion. Ce sondage ne demande pas d'informations auxiliaire. C'est sur lui que se basent toutes les autres méthodes de sondages.

- **Sondage aléatoire simple à probabilité égale sans remise : PESR** : on note $E(\Omega)$: l'ensemble des échantillons aléatoires réalisable et e la réalisation de E . Un plan de taille fixe

$n \in 1, \dots, N$ est dit sans remise si pour tout $e \in E(\Omega)$ on a :

$$p(e = E) = \frac{1}{C_N^n}$$

$\pi_i, \pi_{i,j}$ les probabilités d'inclusion respectivement d'ordre 1 et 2 valent :

$\pi_i = \frac{n}{N}$ et $\pi_{i,j} = \frac{n(n-1)}{N(N-1)}$ pour tout $i, j \in P$ avec $i \neq j$. Le π -estimateur de la moyenne est donné par l'expression :

$$\hat{Y}_{pesr} = \frac{1}{N} \sum_{i \in E} y_i = \bar{y}$$

celui du total T est donné par :

$$\hat{T}_\pi = \frac{N}{n} \sum_{i \in E} y_i$$

. La variance de l'estimateur de la moyenne est donnée par l'expression :

$$Var(\hat{Y}_{pesr}) = (1 - f) \frac{\sigma_c^2}{n}$$

où σ_c^2 et f désigne respectivement la variance corrigée et le taux de sondage. On estime la variance de l'estimateur de la moyenne par :

$$\hat{Var}(\hat{Y}_{pesr}) = (1 - f) \frac{s_c^2}{n}$$

avec

$$s_c^2 = \frac{1}{n-1} \sum_{i \in E} y_i$$

- **Sondage aléatoire simple à probabilité égale avec remise : PEAR** : Un plan de taille fixe $n \in 1, \dots, N$ est dit avec remise si pour tout $e \in E(\Omega)$ on a :

$$p(E = e) = \frac{1}{N^n}$$

On effectue d'une manière répétitive n tirages identitiques et indépendants. Tous les y_i sont iid².

Le π -estimateur de la moyenne de l'échantillon vaut \bar{y} et l'estimateur de sa variance est :

$$\hat{Var}(\hat{Y}_{pear}) = \frac{s_c^2}{n}$$

Sondage à probabilités inégales

Dans le cas où certaines informations supplémentaires sont connues, on peut utiliser autre plan de sondage que le SAS : le plan de sondage à probabilité. Cette méthode est d'autant plus intéressante qu'elle permet d'introduire de l'information auxiliaire dans le plan de sondage ; ainsi la selection se fera avec des probabilités inégales.

Ainsi il est primordial de disposer d'une informations auxiliaire afin d'utiliser cette méthode. Soient

2. indépendant et identiquement distribuée

$X_i, i \in P$ les valeurs prises par la variable auxiliaire. On veut sélectionner n unités d'observation avec des probabilités proportionnelles aux X_i .

Calculons donc les π_i on a : $\pi_i = \frac{nX_i}{\sum_{i \in P} X_i}$. A noter que $\sum_{i \in P} \pi_i = n$.

- **Avec remise : Méthode de Hasen et Hurwitz** $\pi_i = \frac{X_i}{\sum_{i \in P} X_i}$ et $\nu_i = \sum_{j=1}^i p_i$ avec $\nu_0 = 0$.

Algorithme : On génère une variable aléatoire continue u dans $[0, 1]$ et on sélectionne l'unité i tel que : $\nu_{i-1} \leq u < \nu_i$. L'opération sera répéter n fois de manière indépendante pour obtenir l'échantillon de taille n . L'estimateur de la moyenne est :

$$\hat{Y}_{piar} = \frac{1}{nN} \sum_{i \in E} \frac{y_i}{p_i}$$

et sa variance vaut :

$$Var(\hat{Y}_{piar}) = \frac{1}{nN^2} \sum_{i \in P} \left(\frac{Y_i}{p_i} - T \right)^2 p_i$$

- **Sans remise :** Il y a plusieurs méthodes de tirages sans remise proportionnellement à une variable auxiliaire. En effet chaque tirage entraîne la modification des conditions des tirages qui suivent.

La stratification

Cette dernière intervient dans le but de régler combler les limites des SAS (parmi lesquelles : la non-représentation ou la sur-représentation de certaines catégories), mais aussi améliorer la précision. Elle introduit des variables auxiliaires. Elle consiste à partitionner la population P en M sous-populations appelée strates. Puis dans chaque strate h on procède à un SAS d'une manière indépendante.

$P = \cup_{h=1}^M P_h$; $P_h \cap P_k = \emptyset$ pour tout $h \neq k$; $card(P_h) = N_h$: effectif de la strate h .

On a :

$$\sum_{h=1}^M N_h = N$$

La moyenne dans la strate h vaut :

$$\bar{Y}_{strat} = \frac{1}{N_h} \sum_{i \in P_h} y_i$$

L'estimateur de la moyenne dans une strate vaut :

$$\hat{Y}_{strat} = \sum_{h=1}^M \frac{N_h}{N} \bar{y}_h$$

avec :

$$\bar{y}_h = \frac{1}{n_h} \sum_{i \in E_h} y_i$$

La variance de l'estimateur de la moyenne dans une strate est :

$$Var(\hat{Y}_{strat}) = \sum_{h=1}^M \frac{N_h^2}{N^2} \frac{1}{n_h} (1 - f_h) \sigma_{hc}^2$$

avec $f_h = \frac{n_h}{N_h}$ et $\sigma_{hc}^2 = \frac{1}{n_h-1} \sum_{i \in E_h} (y_i - \bar{y})^2$, pour $h = 1, \dots, M$.

– **Plan stratifié à allocations proportionnelles :**

Si pour $h \in 1, \dots, M$, $f_h = \frac{n_h}{N_h} = \frac{n}{N} = f$ avec n fixe, alors on parle de plan stratifié à allocations proportionnelles.

L'estimateur de la moyenne dans une strate est donné par :

$$\sum_{h=1}^M \frac{N_h}{N} \frac{1}{n_h} \sum_{i \in E_h} y_i = \frac{1}{n} \sum_{i_h} y_h$$

et sa variance vaut :

$$(1 - f) \frac{\sigma_c^2}{n}$$

– **Plan stratifié à allocation optimales :**

Elle intervient pour optimiser la précision globale. En effet l'allocation proportionnelle améliore la précision sans l'optimiser (l'homogénéité dans la strate non prise en compte).

Sondage par grappes

Sondage qui consiste à diviser la population en petits groupes, et au fur et à mesure de la sélection, on ne retient que les candidats de chaque groupe correspondant au profil recherché.

Chapitre 2

Fiche d'enquête

2.1 Période d'enquête

L'enquête s'est déroulée du 20 novembre 2022 au 15 décembre 2022.

2.2 Enquêteurs

Les enquêteurs sont :

- **BACCOUCH Ismail**
- **KHORCHANI Amal**
- **SBAA Syrine**

2.3 Technique d'enquête

Les enquêtes se sont déroulées en face à face, dans lrs différentes salles du sport suivantes :

- Fitness Factory - Fouchena
- Star Fitness Gym - El Aouina
- Doctor Gym - La Soukra

2.4 Traitement de données

Les données ont été en codées avec **EXCEL** et traitées avec le logiciel statistique **R**.

2.5 Population source

La population tunisienne.

2.6 Population cible

Les tunisiens qui pratiquent du sport.

2.7 Variable d'intérêt et sa signification

La variable d'intérêt choisie est le budget mensuel alloué à la pratique du sport.

2.8 Collecte des données

Pour la collecte des données, un échantillon de 100 individus choisi d'une manière aléatoire dans le gouvernorat de l'Arianna puisque cette dernière reflète totalement les caractéristiques de la population tunisienne. Ils ont répondu au questionnaire suivant :

2.9 Questionnaire

Le questionnaire a été rédigé avec la supervision de Mme Ouaili Héra.

1. Salle du sport

2. Êtes-vous ?

- Femme
- Homme

3. Quel est votre âge ?

- Entre 15 et 24ans
- Entre 25 et 34ans
- Plus de 35ans

4. Quelle est votre categorie socio-professionnelle ?

- Étudiant(e)/Élève
- Dans la vie active/ à la recherche d'emploi
- À la retraite/ inactif(ve)

5. Quel est votre statut matrimonial ?

- Célibataire
- Marié(e)
- Divorcé(e)/veuf(ve)

6. Si oui , parmi les activités sportives suivantes, lesquelles pratiquez-vous de manière régulière ?

- Fitness, Musculation,Gymnastique
- Running , course à pieds
- Sport collectif (football, basketball, volleyball, rugby, handball...)
- Sport de combat et arts martiaux(boxe, lutte, escrime,karaté, judo...)
- Sport nautique (natation, aquagym,aquabike)
- Autre

7. Combien d'heures par semaine estimez-vous passer dans une salle de sport ?

- Entre 1 et 3 heures
- Entre 3 et 5 heures

- Entre 5 et 7 heures
- Plus de 7 heures

8. Vous pratiquez votre sport principalement ?

- En salle de sport
- À la maison
- En plein air

9. En dehors de l'abonnement de sport , avez-vous d'autres dépenses pour le sport ?

- Oui
- Non

10. Quel est votre budget mensuel estimé pour le sport ?

- Moins 50 Dinars
- Entre 50 et 100 Dinars
- Plus de 100 Dinars

2.10 Base des données

Salle du Sport	Sexe	Age	Categorie socio-professionnelle	Statut Matrimonial	Fitness, Musculation,Gymnastique	Running , course à pieds	Sport collectif
Doctor Gym - Soukra	Femme	Entre 15 et 24 ans	Etudiant(e)/Elève	Célibataire	Fitness, Musculation,Gymnastique	Running , course à pieds	
Doctor Gym - Soukra	Femme	Entre 15 et 24 ans	Etudiant(e)/Elève	Célibataire	Fitness, Musculation,Gymnastique		
Fitness Factory - Fouchana	Homme	Entre 25 et 34 ans	Dans la vie active/ à la recherche d'emploi	Célibataire	Fitness, Musculation,Gymnastique		
Doctor Gym - Soukra	Homme	Entre 25 et 34 ans	Dans la vie active/ à la recherche d'emploi	Célibataire			Sport collectif (football, basketball, volleyball, rugby, handball...)
Fitness Factory - Fouchana	Homme	Entre 15 et 24 ans	Etudiant(e)/Elève	Célibataire			Sport collectif (football, basketball, volleyball, rugby, handball...)
Doctor Gym - Soukra	Homme	Entre 15 et 24 ans	Etudiant(e)/Elève	Célibataire	Fitness, Musculation,Gymnastique		
Doctor Gym - Soukra	Homme	Plus de 35 ans	Dans la vie active/ à la recherche d'emploi	Célibataire	Fitness, Musculation,Gymnastique		
Doctor Gym - Soukra	Femme	Entre 25 et 34 ans	Dans la vie active/ à la recherche d'emploi	Célibataire	Fitness, Musculation,Gymnastique		
Star Fitness Gym - El Aouina	Femme	Entre 25 et 34 ans	Etudiant(e)/Elève	Marî(e)	Fitness, Musculation,Gymnastique		
Fitness Factory - Fouchana	Femme	Entre 15 et 24 ans	Etudiant(e)/Elève	Célibataire	Fitness, Musculation,Gymnastique	Running , course à pieds	
Doctor Gym - Soukra	Homme	Entre 25 et 34 ans	Dans la vie active/ à la recherche d'emploi	Célibataire	Fitness, Musculation,Gymnastique		
Fitness Factory - Fouchana	Femme	Plus de 35 ans	Dans la vie active/ à la recherche d'emploi	Marî(e)	Fitness, Musculation,Gymnastique		
Star Fitness Gym - El Aouina	Femme	Entre 25 et 34 ans	Dans la vie active/ à la recherche d'emploi	Célibataire	Fitness, Musculation,Gymnastique		
Star Fitness Gym - El Aouina	Homme	Entre 15 et 24 ans	Etudiant(e)/Elève	Célibataire	Fitness, Musculation,Gymnastique		Sport collectif (football, basketball, volleyball, rugby, handball...)
Star Fitness Gym - El Aouina	Homme	Entre 15 et 24 ans	Etudiant(e)/Elève	Célibataire	Fitness, Musculation,Gymnastique		Sport collectif (football, basketball, volleyball, rugby, handball...)
Star Fitness Gym - El Aouina	Homme	Entre 25 et 34 ans	Dans la vie active/ à la recherche d'emploi	Marî(e)	Fitness, Musculation,Gymnastique		Sport collectif (football, basketball, volleyball, rugby, handball...)
Doctor Gym - Soukra	Femme	Entre 25 et 34 ans	Etudiant(e)/Elève	Célibataire	Fitness, Musculation,Gymnastique	Running , course à pieds	
Star Fitness Gym - El Aouina	Femme	Entre 15 et 24 ans	Etudiant(e)/Elève	Célibataire	Fitness, Musculation,Gymnastique	Running , course à pieds	
Star Fitness Gym - El Aouina	Femme	Entre 25 et 34 ans	Dans la vie active/ à la recherche d'emploi	Célibataire	Fitness, Musculation,Gymnastique		
Doctor Gym - Soukra	Femme	Plus de 35 ans	Dans la vie active/ à la recherche d'emploi	Marî(e)		Running , course à pieds	
Fitness Factory - Fouchana	Femme	Entre 25 et 34 ans	Dans la vie active/ à la recherche d'emploi	Célibataire	Fitness, Musculation,Gymnastique		
Doctor Gym - Soukra	Homme	Entre 15 et 24 ans	Etudiant(e)/Elève	Célibataire	Fitness, Musculation,Gymnastique		
Doctor Gym - Soukra	Homme	Entre 25 et 34 ans	Dans la vie active/ à la recherche d'emploi	Célibataire	Fitness, Musculation,Gymnastique		Sport collectif (football, basketball, volleyball, rugby, handball...)
Fitness Factory - Fouchana	Femme	Entre 15 et 24 ans	Dans la vie active/ à la recherche d'emploi	Marî(e)	Fitness, Musculation,Gymnastique		
Doctor Gym - Soukra	Homme	Entre 15 et 24 ans	Etudiant(e)/Elève	Célibataire	Fitness, Musculation,Gymnastique		Sport collectif (football, basketball, volleyball, rugby, handball...)
Star Fitness Gym - El Aouina	Femme	Plus de 35 ans	À la retraite/ inactif(ve)	Marî(e)		Running , course à pieds	
Star Fitness Gym - El Aouina	Femme	Entre 25 et 34 ans	Dans la vie active/ à la recherche d'emploi	Célibataire	Fitness, Musculation,Gymnastique		
Fitness Factory - Fouchana	Homme	Entre 25 et 34 ans	Dans la vie active/ à la recherche d'emploi	Célibataire	Fitness, Musculation,Gymnastique	Running , course à pieds	
Star Fitness Gym - El Aouina	Homme	Entre 15 et 24 ans	Etudiant(e)/Elève	Célibataire	Fitness, Musculation,Gymnastique	Running , course à pieds	Sport collectif (football, basketball, volleyball, rugby, handball...)
Fitness Factory - Fouchana	Homme	Entre 15 et 24 ans	Etudiant(e)/Elève	Célibataire	Fitness, Musculation,Gymnastique	Running , course à pieds	Sport collectif (football, basketball, volleyball, rugby, handball...)
Fitness Factory - Fouchana	Femme	Plus de 35 ans	Dans la vie active/ à la recherche d'emploi	Célibataire	Fitness, Musculation,Gymnastique	Running , course à pieds	
Star Fitness Gym - El Aouina	Homme	Entre 25 et 34 ans	Dans la vie active/ à la recherche d'emploi	Marî(e)	Fitness, Musculation,Gymnastique	Running , course à pieds	Sport collectif (football, basketball, volleyball, rugby, handball...)

FIGURE 2.1 – Base des données

Chapitre 3

Resultats

3.1 Analyse des statistiques descriptives

3.1.1 Sexe

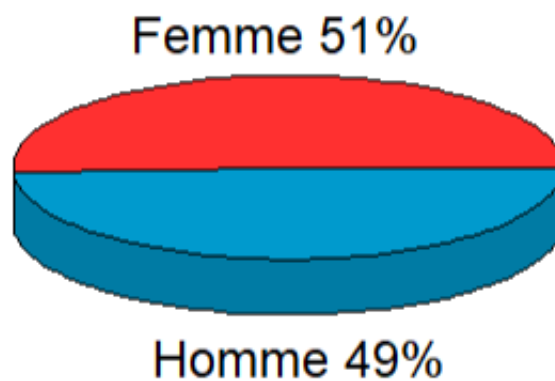


FIGURE 3.1 – Répartition des individus selon le sexe

D'après la figure au-dessus, on constate qu'il existe plus que femmes que des hommes dans notre population cible. Autrement dit, selon cette population les femmes pratiquent du sport plus que les hommes.

3.1.2 Age

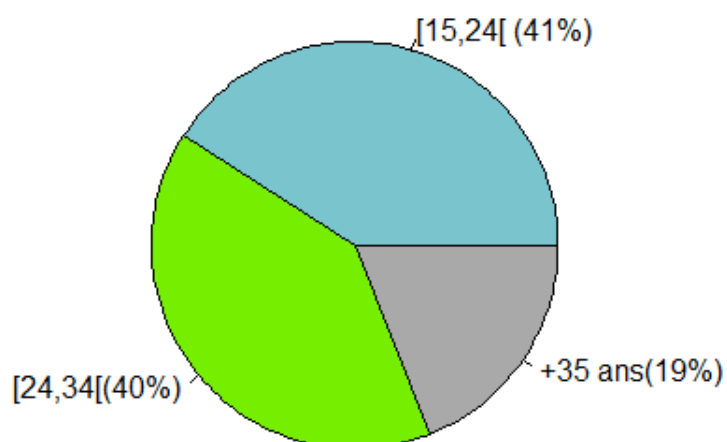


FIGURE 3.2 – Répartition des individus selon l'âge

Selon la figure au-dessus, on peut dire que la catégorie des individus qui ont l'âge de 15ans à 24ans est plus représentée que les autres catégories.

3.1.3 Catégorie socio-professionnelle

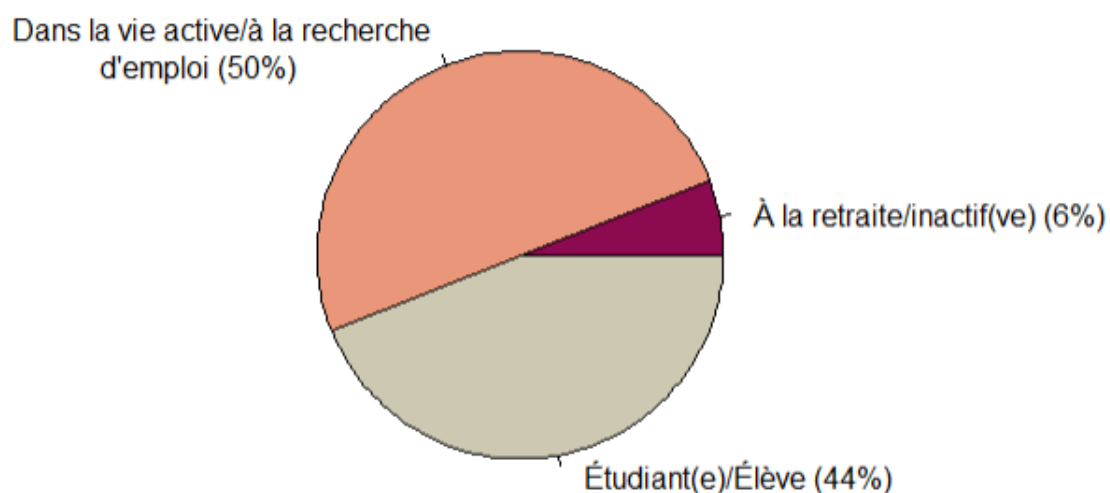


FIGURE 3.3 – Répartition des individus selon leur catégorie socio-professionnelle

D'après cette figure, on constate que les individus qui sont dans la vie active ou à la recherche d'emploi sont plus nombreux que les individus qui sont des étudiants ou des élèves et à ceux qui sont à la retraite ou inactive.

3.1.4 Statut matrimonial

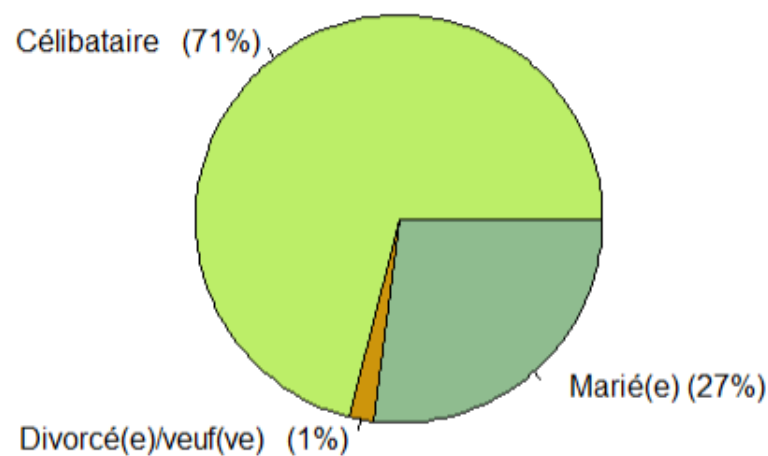


FIGURE 3.4 – Répartition des individus selon leur statut matrimonial

En constatant la figure au-dessus, on peut dire que la catégories des individus qui sont célibataire est plus représentée que d'autres catégories.

3.1.5 Activités sportives pratiquées

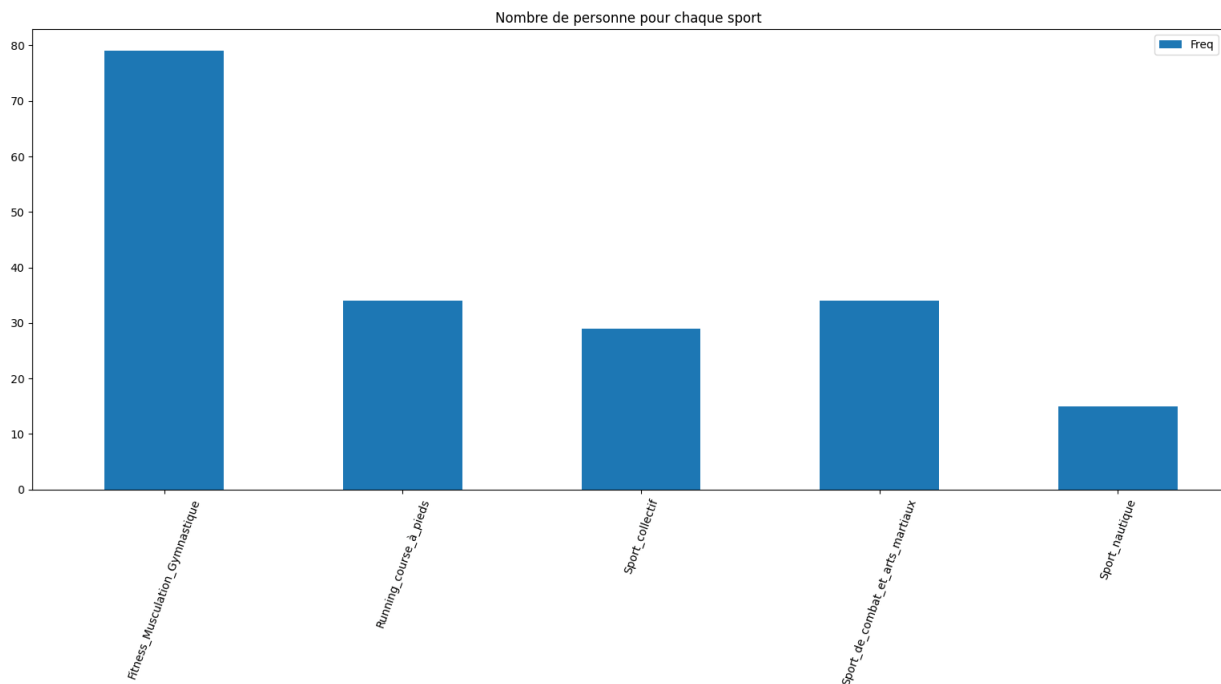


FIGURE 3.5 – Répartition des individus selon leurs activités sportives

Selon la figure au-dessus, on remarque que la **Fitness, Musculation, Gymnastique** est l'activité la plus pratiquée parmi les activités sportives proposées. Alors que le **Sport nautique** est le moins pratiquées.

3.1.6 Heures

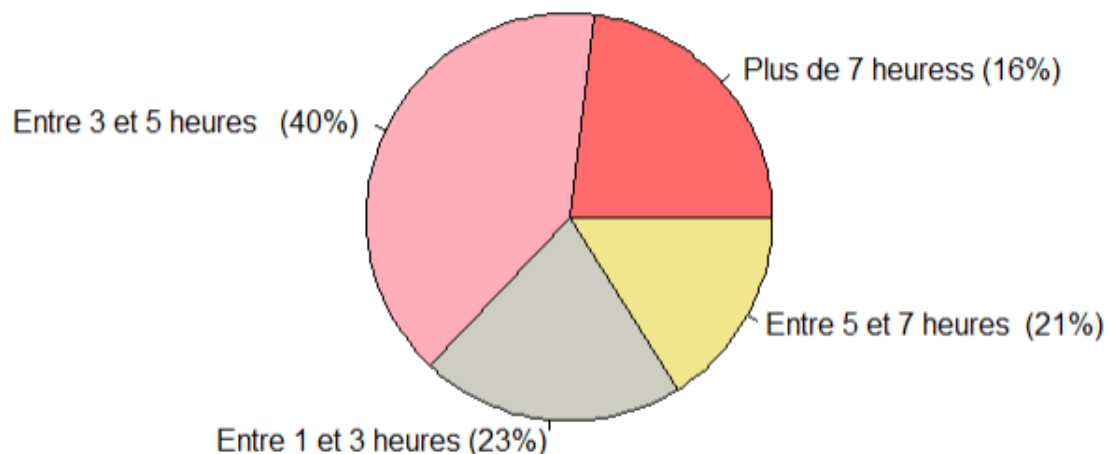


FIGURE 3.6 – Répartition des individus selon les heures passées

D'après la figure au-dessus, on note que 40 parmi 100 individus passent entre 3 et 5 heures par semaine pratiquant du sport. Et que seulement 16 individus passent plus de 7 heures par semaine pratiquant du sport.

3.1.7 Endroit

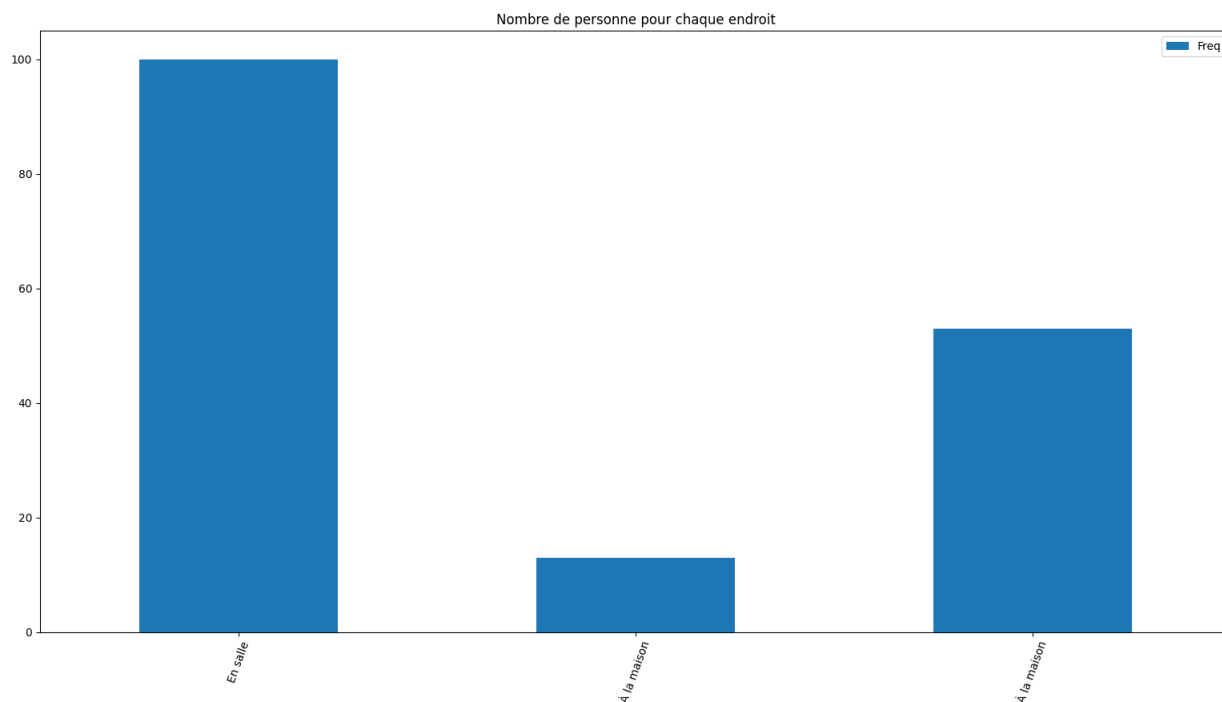


FIGURE 3.7 – Répartition des individus selon leurs endroits

Selon la figure au-dessus, on constate que la majorité de notre population cible pratique le sport principalement en salle du sport. Alors que la minorité des individus pratiquent le sport à la maison.

3.1.8 Autres dépenses

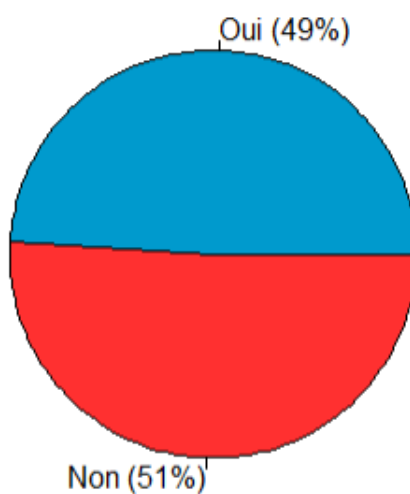


FIGURE 3.8 – Répartition des individus selon leurs dépenses supplémentaires

D'après la figure au-dessus, on remarque que 51 parmi 100 individus ont d'autres dépenses pour le sport en dehors de l'abonnement de sport.

3.1.9 Budget mensuel alloué à la pratique sport

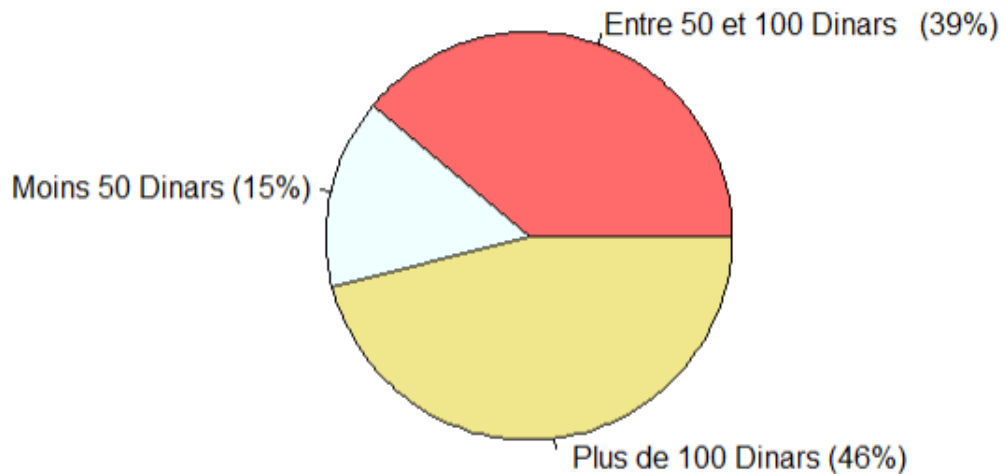


FIGURE 3.9 – Répartition des individus selon leur budget mensuel

Dans la figure au-dessus, on constate que 46 parmi 100 individus allouent plus de 100 Dinars à la pratique du sport.

3.2 Traitement des données

On a $N = 100$ individus. En calculant la moyenne et la variance du budget alloué à la pratique du sport, on obtient $\bar{Y} = 103.25$ et $\sigma^2 = 2177.967$. On choisira $n = 20$. Pour chaque méthodes, on effectuera l'opération de selection de l'échantillon et d'estimation ponctuelle et par intervalle de confiance 20 fois.

Ensuite on représentera les 20 intervalles de confiance obtenus.

3.2.1 Sondage aléatoire simple :le SAS

Sondage aléatoire simple à probabilité égale sans remise : le PESR

Les budgets alloués à la pratique du sport dans les 20 échantillons tirés :

```
> round(Mt_ech,1) # les 20 échantillons sont tirés et voici les résultats
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12] [,13] [,14] [,15] [,16] [,17] [,18] [,19] [,20]
[1,] 200 180 150 90 150 100 180 200 100 65 75 100 150 65 100 100 100 150 80 150
[2,] 110 65 70 75 65 65 120 40 70 120 150 80 170 170 150 180 120 110 65 40
[3,] 65 40 200 120 170 75 130 55 200 90 40 75 70 40 40 130 40 150 150 70
[4,] 250 90 40 55 70 130 150 150 75 130 70 40 65 150 40 90 90 200 90 150
[5,] 150 150 170 250 150 150 160 65 170 65 120 120 110 65 55 130 130 130 70 65
[6,] 120 40 130 190 200 40 90 150 40 120 90 150 65 40 150 80 40 150 130 120
[7,] 40 150 150 70 150 65 150 130 150 130 130 160 190 150 110 90 120 190 65 150
[8,] 65 250 110 90 40 130 120 80 40 120 90 65 130 190 130 70 90 130 120 110
[9,] 150 150 65 40 120 90 90 120 65 120 40 110 65 65 65 40 70 65 65 65
[10,] 80 130 150 40 40 70 40 150 120 180 120 65 150 120 90 65 120 40 150 120
[11,] 90 150 190 120 65 40 40 80 70 65 65 120 80 150 40 60 40 120 80 40
[12,] 70 40 40 60 90 120 180 150 90 130 65 150 90 80 120 40 65 65 120 40
[13,] 90 65 120 40 150 150 65 65 180 90 150 90 40 65 40 65 60 130 90 40
[14,] 90 40 90 65 140 180 40 70 180 90 120 40 90 65 40 180 150 65 60 65
[15,] 140 65 120 150 120 90 140 120 90 130 65 80 120 150 120 40 65 40 40 130
[16,] 120 65 40 70 90 140 40 130 120 120 40 120 65 120 60 120 120 120 65 120
[17,] 70 120 40 120 65 90 40 65 120 65 80 120 120 120 150 90 65 120 140 180
[18,] 90 130 90 65 90 120 120 40 90 140 120 120 40 65 40 65 65 130 80 140
[19,] 120 65 120 40 200 40 40 65 65 65 40 40 120 65 70 40 120 120 120 90
[20,] 150 150 80 120 150 150 65 40 40 40 150 150 40 80 65 40 200 140 40 120
```

FIGURE 3.10 – **Echantillons PESR**

Les estimateurs de la moyenne de l'IMC dans les 20 échantillons tirés :

113.00	106.75	108.25	93.50	115.75	101.75	100.00	98.25	103.75	103.75
91.00	99.75	98.50	100.75	83.75	85.75	93.50	118.25	91.00	100.25

Les estimateurs de la variance dans les 20 échantillons tirés :

2535.263	3332.303	2550.724	2997.632	2469.145	1742.829	2597.368	2208.618	2478.618	1228.618
1504.211	1472.303	1966.053	2145.461	1720.724	1884.934	1734.474	1932.303	1214.737	1956.513

Les 20 intervalles de confiance :

108.96319	117.03681
102.71319	110.78681
104.21319	112.28681
89.46319	97.53681
111.71319	119.78681
97.71319	105.78681
95.96319	104.03681
94.21319	102.28681
99.71319	107.78681
99.71319	107.78681
86.96319	95.03681
95.71319	103.78681
94.46319	102.53681
96.71319	104.78681
79.71319	87.78681
81.71319	89.78681
89.46319	97.53681
114.21319	122.28681
86.96319	95.03681
96.21319	104.28681

Les longueurs d'intervalles de confiance :

8.073613 8.073613 8.073613 8.073613 8.073613 8.073613 8.073613 8.073613 8.073613 8.073613
8.073613 8.073613 8.073613 8.073613 8.073613 8.073613 8.073613 8.073613 8.073613 8.073613

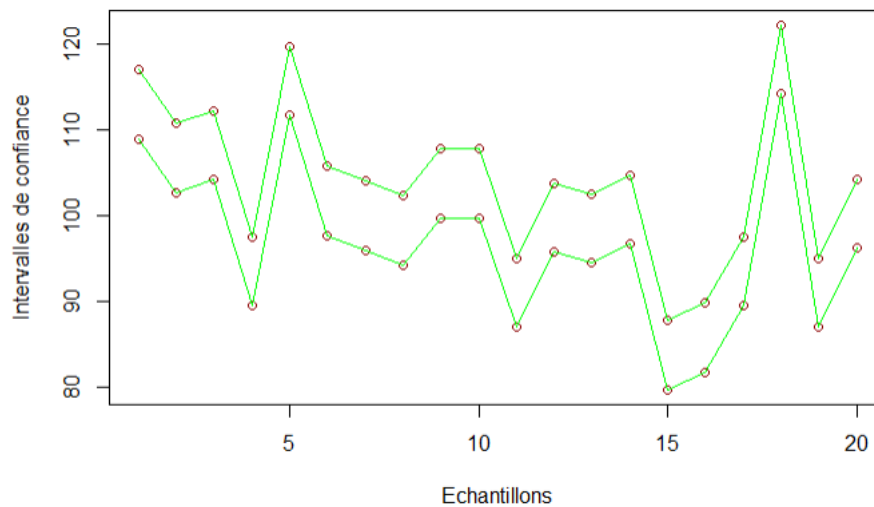


FIGURE 3.11 – **Intervalle de confiance pour PESR**

La moyenne des 20 estimateurs de la moyenne est :**100.3625**.

Sondage aléatoire simple à probabilité égale avec remise : le PEAR

Les budgets alloué à la pratique du sport dans les 20 échantillons tirés :

```
> round(ech1,1) #les 20 échantillons sont tirés et les résultats
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12] [,13] [,14] [,15] [,16] [,17] [,18] [,19] [,20]
[1,] 130 200 120 200 130 80 180 120 65 200 40 40 65 65 65 180 90 180 40 130
[2,] 120 90 200 90 80 65 70 90 40 150 140 120 90 90 130 130 90 40 90 90
[3,] 120 120 65 40 140 150 40 40 40 110 70 120 80 60 130 150 150 70 65 65
[4,] 120 40 150 120 90 55 60 120 65 80 70 40 65 55 70 40 150 160 120 40
[5,] 120 150 130 200 140 110 120 130 70 55 65 120 40 70 150 120 190 150 65 200
[6,] 120 140 65 90 80 120 120 120 180 90 100 90 170 90 75 130 90 100 120 70
[7,] 120 65 65 65 40 40 60 40 90 40 90 90 90 120 90 80 140 120 120 150
[8,] 65 120 200 120 65 65 55 80 140 160 65 120 70 90 40 120 40 120 120 140
[9,] 90 120 65 65 40 130 140 40 140 150 90 65 65 120 65 80 40 120 60 200
[10,] 65 70 120 120 40 75 90 40 40 70 180 90 150 40 150 200 150 120 55 140
[11,] 200 65 40 90 40 90 120 190 120 120 80 90 90 40 40 90 90 180 120 120
[12,] 180 150 140 120 120 130 150 70 90 40 80 40 40 130 70 130 65 150 130 80
[13,] 65 40 40 130 80 140 90 160 180 140 130 80 140 120 40 40 150 90 180 150
[14,] 120 100 40 70 150 90 180 100 40 130 130 40 75 150 150 40 90 180 40 40
[15,] 120 120 150 40 65 80 120 90 120 90 65 160 65 90 80 90 40 190 120 120
[16,] 65 110 90 65 40 40 120 90 90 40 90 150 65 40 65 40 90 160 90 40
[17,] 65 70 150 130 80 90 70 140 90 150 150 110 90 90 120 70 180 200 40 110
[18,] 120 90 80 110 200 90 40 90 180 40 40 120 75 180 120 90 65 65 90 65
[19,] 65 110 170 80 90 130 120 170 65 150 170 150 65 40 120 190 130 65 90 130
[20,] 90 70 65 140 70 100 150 40 40 140 150 65 150 90 140 80 150 90 140 200
```

FIGURE 3.12 – **Echantillon PEAR**

Les estimateurs de la moyenne de l'IMC dans les 20 échantillons tirés :

108.00	102.00	107.25	104.25	89.00	93.50	104.75	98.00	94.25	107.25
99.75	95.00	87.00	88.50	95.50	104.50	109.00	127.50	94.75	114.00

Les estimateurs de la variance dans les 20 échantillons tirés :

1430.000	1624.737	2740.724	1958.618	1985.789	1050.263	1882.829	2101.053	2395.461	2372.303
1761.776	1465.789	1348.421	1531.842	1531.316	2468.158	2196.316	2269.737	1474.934	2717.368

Les 20 intervalles de confiance :

103.71825	112.28175
97.71825	106.28175
102.96825	111.53175
99.96825	108.53175
84.71825	93.28175
89.21825	97.78175
100.46825	109.03175
93.71825	102.28175
89.96825	98.53175
102.96825	111.53175
95.46825	104.03175
90.71825	99.28175
82.71825	91.28175
84.21825	92.78175
91.21825	99.78175
100.21825	108.78175
104.71825	113.28175
123.21825	131.78175
90.46825	99.03175
109.71825	118.28175

Les longueurs des intervalles de confiance :

8.563498	8.563498	8.563498	8.563498	8.563498	8.563498	8.563498	8.563498	8.563498	8.563498
8.563498	8.563498	8.563498	8.563498	8.563498	8.563498	8.563498	8.563498	8.563498	8.563498

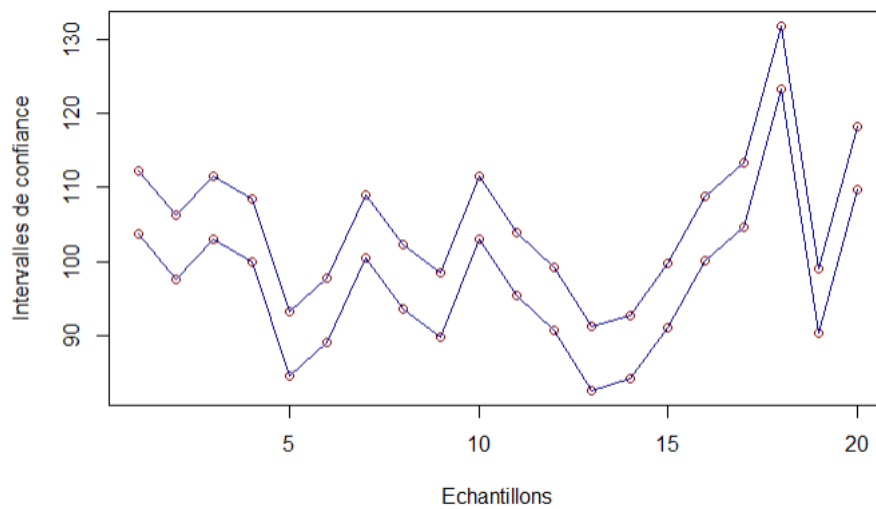


FIGURE 3.13 – Intervalles de confiance pour PEAR

La moyenne des 20 estimateurs de la moyenne est : **101.1875**

3.2.2 Sondage aléatoire à probabilité inégales

Sans remise : le PISR

Le budget mensuel dans les 20 échantillons tirés :

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]	[,10]	[,11]	[,12]	[,13]	[,14]	[,15]	[,16]	[,17]	[,18]	[,19]	[,20]
[1,]	90	100	90	180	100	100	100	100	180	100	100	180	100	100	80	80	100	100	90	100
[2,]	65	150	65	200	70	150	70	150	200	70	150	200	150	150	200	200	150	150	65	150
[3,]	150	120	150	120	40	120	120	120	120	120	120	120	120	75	150	150	120	120	150	75
[4,]	110	170	110	180	180	170	180	170	180	180	170	180	170	170	90	90	170	170	110	170
[5,]	160	70	160	150	130	130	130	70	150	130	130	150	70	70	160	150	130	130	160	70
[6,]	150	55	150	120	65	55	65	55	120	65	55	120	55	40	150	150	55	55	200	40
[7,]	150	200	150	250	250	200	250	200	250	250	200	250	200	200	150	250	200	200	150	200
[8,]	90	110	90	65	130	110	130	110	65	130	110	130	110	110	120	120	110	110	90	110
[9,]	190	65	190	190	150	65	150	65	150	150	65	150	65	40	190	190	65	65	130	40
[10,]	65	120	65	130	120	120	120	130	130	120	120	130	120	130	80	80	120	120	65	130
[11,]	150	90	150	120	90	90	90	120	70	90	90	70	90	120	120	120	90	90	150	120
[12,]	40	90	40	90	120	90	120	90	90	120	90	120	90	90	40	40	90	90	120	90
[13,]	65	40	65	180	80	80	80	40	180	80	80	180	40	120	65	65	80	80	65	120
[14,]	150	90	150	130	130	90	40	90	130	40	90	130	90	90	150	90	90	90	150	90
[15,]	65	140	65	120	140	140	140	65	120	140	140	40	140	65	65	120	140	140	65	65
[16,]	90	150	90	180	180	150	180	150	180	180	150	180	150	150	70	70	150	150	90	150
[17,]	120	120	120	90	140	120	140	120	90	140	120	140	120	120	120	120	120	120	120	120
[18,]	65	130	65	90	120	120	120	130	65	120	120	65	120	130	40	40	120	120	65	130
[19,]	40	140	90	120	80	80	80	140	120	80	80	120	140	140	90	90	80	80	40	140
[20,]	150	120	150	200	200	40	200	120	200	200	40	200	120	120	150	150	40	40	150	120

FIGURE 3.14 – **Echantillon PISR**

Les estimateurs de la moyenne de budget mensuel dans les 20 échantillons tirés :

107.75	113.50	110.25	145.25	125.75	111.00	125.25	111.75	139.50	125.25
111.00	142.75	113.00	111.50	114.00	118.25	111.00	111.00	111.25	111.50

Les estimateurs de la variance dans les 20 échantillons tirés :

2085.461	1584.474	1853.882	2267.039	2540.197	1559.474	2540.724	1653.355	2570.789	2540.724
1559.474	2535.461	1572.105	1710.789	2264.737	2945.461	1559.474	1559.474	1862.829	1710.789

Les 20 intervalles de confiance :

102.6542	112.8458
108.4042	118.5958
105.1542	115.3458
140.1542	150.3458
120.6542	130.8458
105.9042	116.0958
120.1542	130.3458
134.4042	144.5958
120.1542	130.3458
105.9042	116.0958
137.6542	147.8458
107.9042	118.0958
106.4042	116.5958
108.9042	119.0958
113.1542	123.3458
105.9042	116.0958
105.9042	116.0958
106.1542	116.3458
106.4042	116.5958

Les longueurs des intervalles de confiance :

10.19163	10.19163	10.19163	10.19163	10.19163	10.19163	10.19163	10.19163	10.19163	10.19163
10.19163	10.19163	10.19163	10.19163	10.19163	10.19163	10.19163	10.19163	10.19163	10.19163

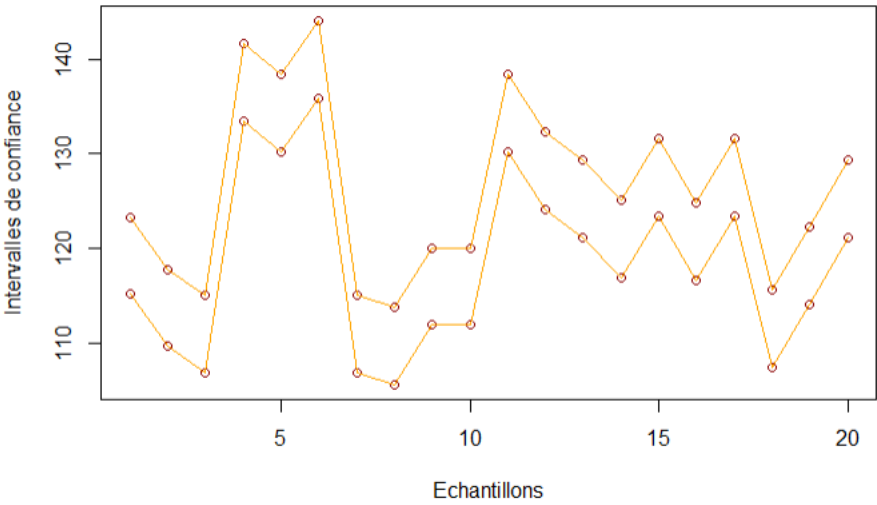


FIGURE 3.15 – Intervalle de confiance pour PISR

La moyenne des 20 estimateurs de la moyenne est : 118.525

Avec remise : le PIAR

Le budget mensuel dans les 20 échantillons tirés :

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]	[,10]	[,11]	[,12]	[,13]	[,14]	[,15]	[,16]	[,17]	[,18]	[,19]	[,20]
[1,]	80	180	100	100	90	80	100	100	100	100	100	100	90	180	80	90	80	180	100	80
[2,]	200	200	150	150	65	65	150	150	150	150	150	150	65	200	200	65	200	70	150	200
[3,]	150	150	120	75	150	150	120	120	120	120	75	120	150	120	150	170	150	120	120	150
[4,]	90	40	180	170	110	90	180	170	170	170	170	170	90	180	90	110	90	180	170	90
[5,]	160	150	130	110	160	160	130	70	70	70	110	70	160	150	150	160	150	130	70	160
[6,]	150	150	65	160	200	150	65	40	40	55	40	40	150	120	150	200	150	120	55	150
[7,]	150	250	250	200	150	150	200	200	200	200	200	200	150	250	250	150	250	250	200	150
[8,]	120	65	130	150	90	120	110	110	110	110	110	110	120	130	120	90	120	130	110	120
[9,]	190	190	150	90	130	190	150	40	40	65	40	40	190	150	190	130	190	150	65	190
[10,]	80	130	120	130	65	65	120	130	130	130	130	130	65	130	80	120	80	130	130	80
[11,]	150	120	90	120	150	150	90	120	120	120	120	120	150	70	120	150	120	40	120	150
[12,]	40	90	120	150	120	40	120	90	90	90	150	90	40	120	40	120	40	120	90	40
[13,]	65	180	80	120	65	65	80	120	120	40	120	120	65	180	65	65	65	180	40	65
[14,]	150	90	40	60	150	150	40	90	90	90	60	90	150	130	90	150	90	130	90	150
[15,]	65	120	140	65	65	65	140	65	65	65	65	65	65	65	40	120	65	120	140	65
[16,]	70	40	150	150	90	70	150	150	150	150	150	150	70	180	40	90	40	180	150	70
[17,]	120	90	140	120	120	120	140	120	120	120	120	120	120	140	120	130	120	140	120	120
[18,]	65	90	120	130	140	65	120	130	130	130	130	130	65	65	90	140	90	120	130	40
[19,]	90	120	80	140	40	90	80	140	140	140	140	140	90	120	90	40	120	65	140	90
[20,]	150	200	200	120	150	150	40	120	120	120	120	120	150	200	150	150	150	200	120	150

FIGURE 3.16 – **Echantillon PIAR**

Les estimateurs de la moyenne du budget mensuel dans les 20 échantillons tirés :

116.75	132.25	127.75	125.50	115.00	109.25	116.25	113.75	113.75	111.75
115.00	113.75	109.75	142.75	119.25	119.25	120.75	138.75	111.75	115.50

Les estimateurs de la variance dans les 20 échantillons tirés :

2161.250	3211.776	2280.197	1270.789	1778.947	1948.092	1775.987	1639.145	1639.145	1653.355
1776.316	1639.145	1922.303	2535.461	2863.882	1734.934	2816.513	2352.303	1653.355	2328.684

Les 20 intervalles de confiance :

112.637	120.863
128.137	136.363
123.637	131.863
121.387	129.613
110.887	119.113
105.137	113.363
112.137	120.363
109.637	117.863
109.637	117.863
107.637	115.863
110.887	119.113
109.637	117.863
105.637	113.863
138.637	146.863
115.137	123.363
115.137	123.363
116.637	124.863
134.637	142.863
107.637	115.863
111.387	119.613

Les longueurs des intervalles de confiance :

8.226072	8.226072	8.226072	8.226072	8.226072	8.226072	8.226072	8.226072	8.226072	8.226072
8.226072	8.226072	8.226072	8.226072	8.226072	8.226072	8.226072	8.226072	8.226072	8.226072

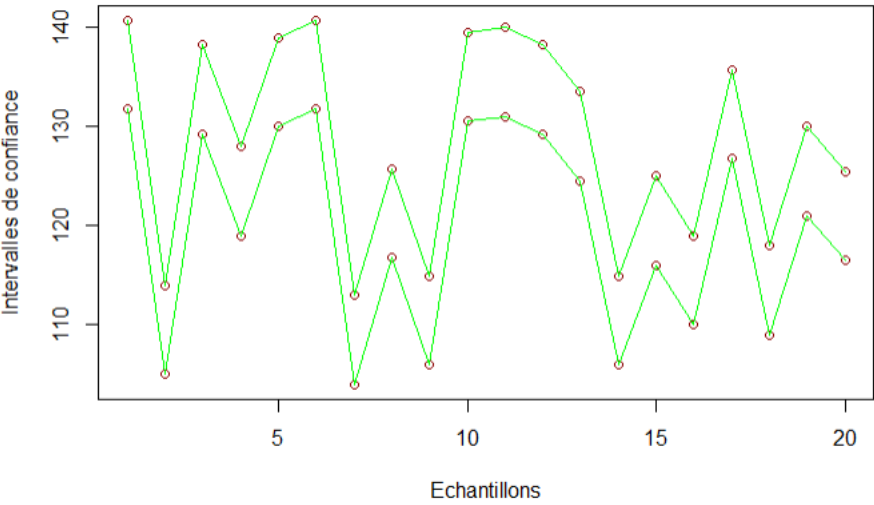


FIGURE 3.17 – Intervalle de confiance pour PIAR

La moyenne des 20 estimateurs de la moyenne est : **124.2**

3.2.3 La stratification

Dans cette partie, il faut choisir une variable aléatoire pour les différentes strates suivantes. Cette variable va être la variable **sexe**.

Strates de même taille

Les budgets mensuels alloués à la pratique du sport dans les 20 échantillons tirés :

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]	[,10]	[,11]	[,12]	[,13]	[,14]	[,15]	[,16]	[,17]	[,18]	[,19]	[,20]
[1,]	200	100	40	65	100	200	150	80	100	100	65	75	170	40	100	100	90	100	200	90
[2,]	40	150	90	180	150	40	75	200	180	65	160	160	180	40	100	120	65	90	170	120
[3,]	180	150	110	130	120	180	150	75	90	120	150	55	150	55	70	170	40	75	90	110
[4,]	40	110	150	40	150	110	110	120	120	40	65	150	55	120	40	70	90	40	200	150
[5,]	130	40	110	55	130	160	55	160	40	40	150	250	150	200	40	160	150	90	130	55
[6,]	65	150	190	65	40	130	110	65	55	160	130	40	40	120	120	40	200	150	65	120
[7,]	120	120	130	250	130	130	80	120	200	40	120	130	130	90	200	130	65	40	90	250
[8,]	65	130	80	150	90	65	65	150	110	65	65	90	120	40	80	65	40	200	40	110
[9,]	190	120	70	130	40	40	120	120	40	65	120	40	90	65	90	150	120	120	190	150
[10,]	65	120	150	65	150	70	40	70	130	130	40	70	120	120	150	65	120	150	130	190
[11,]	40	90	90	90	90	90	120	90	90	90	120	90	80	90	120	90	90	90	90	120
[12,]	120	80	65	40	120	40	40	40	65	40	120	40	90	120	90	40	40	65	90	60
[13,]	60	60	40	180	40	65	150	60	130	120	40	40	140	40	40	180	40	90	140	90
[14,]	130	90	90	120	80	140	40	40	90	150	40	40	70	40	60	60	40	40	40	130
[15,]	65	40	150	40	180	90	150	40	65	140	65	120	90	90	90	140	90	65	65	65
[16,]	120	90	40	65	90	90	180	40	40	120	40	65	120	40	40	40	90	40	150	120
[17,]	120	150	180	120	65	65	90	90	140	120	70	40	120	140	65	120	130	150	120	90
[18,]	120	150	120	90	180	90	80	140	90	40	90	65	65	120	140	65	120	70	130	120
[19,]	65	120	65	120	140	65	120	90	80	80	120	90	120	130	40	140	40	65	65	65
[20,]	120	140	140	150	150	150	200	40	40	150	40	140	200	120	150	90	200	120	90	40

FIGURE 3.18 – **Echatillon pour strate de même taille**

Les estimateurs de la moyenne du budget mensuel dans les 20 échantillons tirés :

118.25 109.75 114.00 96.25 102.50 99.00 103.25 100.75 95.50 133.75
100.25 86.75 106.50 100.50 107.25 103.75 109.25 94.25 101.25
99.75

Les estimateurs de la variance :

3390.197 1401.250 2433.158 2639.145 2193.421 2230.526 1248.092 2161.250 2041.842 2891.776
1382.829 1584.934 1781.842 1791.842 1851.250 2033.882 2453.355 2177.039 2183.882 1635.461

Les 20 intervalles de confiance :

113.88978	122.61022
105.38978	114.11022
109.63978	118.36022
91.88978	100.61022
98.13978	106.86022
94.63978	103.36022
98.88978	107.61022
96.38978	105.11022
91.13978	99.86022
129.38978	138.11022
95.88978	104.61022
82.38978	91.11022
102.13978	110.86022
96.13978	104.86022
102.88978	111.61022
99.38978	108.11022
104.88978	113.61022
89.88978	98.61022
96.88978	105.61022
95.38978	104.11022

Les longueur des intervalles de confiance :

8.720435 8.720435 8.720435 8.720435 8.720435 8.720435 8.720435 8.720435 8.720435 8.720435
8.720435 8.720435 8.720435 8.720435 8.720435 8.720435 8.720435 8.720435 8.720435 8.720435

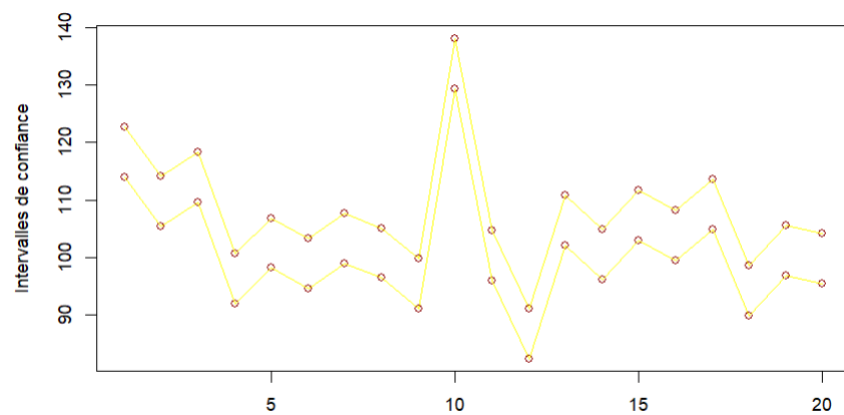


FIGURE 3.19 – – **Intervalle de confiance pour stratification avec strates de même taille**

La moyenne des 20 estimateurs de la moyenne vaut : **104.125**.

Stratification à allocations proportionnelles

Les budgets mensuels dans les 20 échantillons tirés :

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]	[,10]	[,11]	[,12]	[,13]	[,14]	[,15]	[,16]	[,17]	[,18]	[,19]	[,20]
[1,]	100	200	80	100	100	150	80	180	70	70	150	80	180	100	70	120	100	100	120	120
[2,]	150	180	200	90	120	110	150	70	65	120	180	70	40	80	40	40	180	80	40	170
[3,]	70	130	65	75	40	130	75	170	130	120	130	75	110	70	120	170	150	90	120	40
[4,]	200	150	120	150	130	150	120	150	150	150	65	160	130	160	40	40	110	40	150	130
[5,]	70	150	40	90	160	40	40	55	160	170	150	130	200	250	150	130	150	150	90	120
[6,]	40	120	40	150	120	120	120	65	40	70	130	150	90	110	160	160	130	160	65	200
[7,]	130	90	130	65	150	110	150	120	200	200	130	190	150	130	150	150	40	110	130	110
[8,]	65	120	65	120	130	65	190	130	130	110	120	130	130	150	40	250	150	65	40	65
[9,]	120	40	120	40	65	150	130	90	65	65	80	120	80	120	120	150	80	150	120	190
[10,]	150	120	120	150	90	120	130	70	150	120	40	40	70	150	65	120	120	65	150	80
[11,]	90	120	90	120	90	40	120	40	90	90	40	90	40	120	40	80	120	90	90	80
[12,]	80	40	40	60	120	80	80	180	65	120	40	65	40	80	65	180	90	40	40	180
[13,]	180	180	40	90	40	40	65	65	65	120	40	90	80	65	40	130	65	40	65	130
[14,]	150	65	180	40	80	65	65	150	130	130	150	65	90	150	120	150	120	120	150	65
[15,]	40	40	90	90	65	70	120	65	40	120	65	140	65	150	65	180	120	140	140	90
[16,]	65	65	90	140	140	140	70	150	65	65	65	130	140	120	120	40	120	65	130	120
[17,]	40	70	130	40	130	90	120	180	40	180	90	120	40	140	140	70	140	140	120	40
[18,]	70	65	65	65	65	80	120	40	70	130	80	65	65	120	80	120	40	80	65	80
[19,]	120	65	40	180	90	40	40	140	90	90	65	90	65	65	65	120	120	200	40	40
[20,]	40	90	120	200	40	200	200	120	120	150	40	150	130	40	90	90	40	150	200	150

FIGURE 3.20 – **Echantillon pour stratification à allocations proportionnelles**

les estimateurs de la moyenne :

103.50	96.00	103.75	108.50	91.50	115.25	102.00	107.50	100.50	90.75
109.00	109.25	111.00	98.75	98.50	91.75	103.25	98.25	99.00	100.50

Les estimateurs de la variance :

3310.789	1812.105	2041.776	1566.053	1826.579	3430.197	2095.789	3290.789	1528.684	1695.461
1809.474	2395.461	2112.105	1533.882	1939.737	2132.303	2366.513	2537.566	2680.526	2160.263

Les 20 intervalles de confiance :

100.54177	106.45823
93.04177	98.95823
100.79177	106.70823
105.54177	111.45823
88.54177	94.45823
112.29177	118.20823
99.04177	104.95823
104.54177	110.45823
97.54177	103.45823
87.79177	93.70823
106.04177	111.95823
106.29177	112.20823
108.04177	113.95823
95.79177	101.70823
95.54177	101.45823
88.79177	94.70823
100.29177	106.20823
95.29177	101.20823
96.04177	101.95823
97.54177	103.45823

Les longueur des intervalles de confiance :

5.916465 5.916465 5.916465 5.916465 5.916465 5.916465 5.916465 5.916465 5.916465 5.916465
5.916465 5.916465 5.916465 5.916465 5.916465 5.916465 5.916465 5.916465 5.916465 5.916465

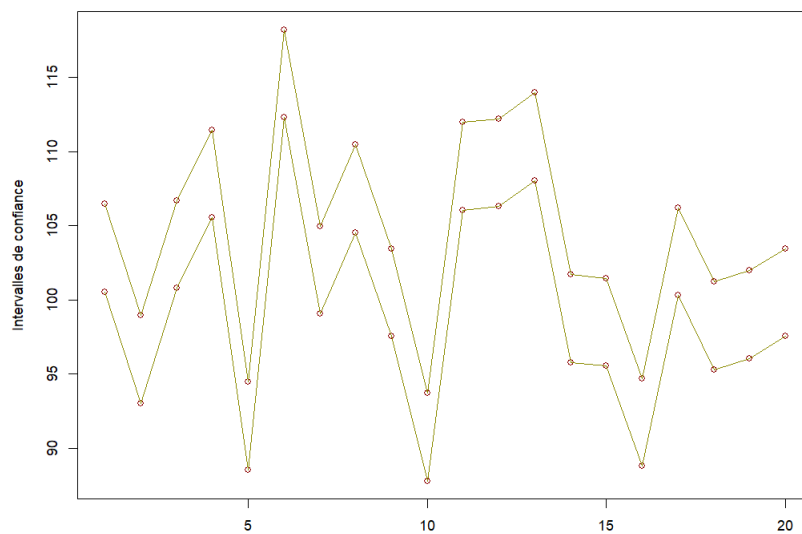


FIGURE 3.21 – **Intervalle de confiance pour stratification à allocations proportionnelles**

La moyenne des 20 estimateurs de la moyenne vaut : **101.925**.

Stratification à allocations optimales

Les budgets mensuels dans les 20 échantillons tirés :

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]	[,10]	[,11]	[,12]	[,13]	[,14]	[,15]	[,16]	[,17]	[,18]	[,19]	[,20]
[1,]	120	100	80	180	80	180	180	100	180	180	90	180	100	75	90	70	90	80	150	150
[2,]	110	90	90	150	90	80	90	90	90	150	150	120	70	40	65	120	75	200	200	40
[3,]	150	40	70	75	150	70	150	150	150	70	150	120	40	150	40	90	150	65	120	160
[4,]	160	180	180	70	40	75	180	65	150	65	40	150	180	150	120	110	110	40	120	40
[5,]	150	70	110	130	150	150	110	170	55	150	55	65	90	130	180	130	200	90	40	150
[6,]	250	160	55	160	170	70	90	90	65	180	65	90	40	120	90	40	150	70	130	130
[7,]	90	250	40	250	130	120	65	70	150	70	130	120	150	130	65	55	120	150	130	130
[8,]	150	110	65	190	150	40	130	110	120	160	65	130	130	40	130	120	40	130	120	130
[9,]	65	65	130	120	250	130	65	90	120	130	120	90	65	120	130	90	190	190	130	120
[10,]	40	120	80	65	65	90	150	40	150	190	65	120	90	150	150	40	130	120	120	150
[11,]	40	40	40	120	180	90	120	40	65	120	90	90	120	90	90	120	90	180	120	120
[12,]	180	120	40	40	130	120	90	40	65	90	120	120	40	40	40	40	90	65	65	80
[13,]	90	40	130	40	120	40	40	150	150	65	180	60	65	120	65	120	65	150	60	65
[14,]	140	180	90	40	150	40	40	40	120	40	65	90	60	80	150	180	40	65	90	120
[15,]	65	65	150	90	40	120	90	90	140	120	40	40	65	90	140	140	90	65	180	65
[16,]	120	40	140	65	70	40	65	120	120	65	40	65	120	90	90	40	65	90	40	120
[17,]	120	40	90	180	90	70	140	140	65	90	120	40	150	150	140	65	40	140	65	120
[18,]	65	120	120	90	130	120	120	80	120	65	40	120	90	150	130	120	140	80	140	65
[19,]	120	120	40	120	140	90	80	120	40	65	65	140	40	70	140	90	65	65	40	90
[20,]	40	40	40	200	40	40	200	40	200	90	200	130	200	120	120	120	150	200	120	140

FIGURE 3.22 – **Echantillon pour stratification à allocations optimales**

Les estimateurs de la moyenne :

96.50 102.50 106.00 112.50 126.25 90.75 108.25 97.00 112.75 99.25
 99.25 99.25 113.75 87.00 107.25 111.25 92.75 101.75 105.75 104.00

Les estimateurs de la moyenne :

1866.053 2372.368 2917.368 2761.842 2736.513 1511.250 1824.408 2177.368 2056.513 1442.829
 2598.092 1313.882 3123.355 1195.789 2585.461 2181.250 1682.829 2221.776 1824.408 1846.316

Les 20 intervalles de confiance :

92.48243	100.51757
98.48243	106.51757
101.98243	110.01757
108.48243	116.51757
122.23243	130.26757
86.73243	94.76757
104.23243	112.26757
92.98243	101.01757
108.73243	116.76757
95.23243	103.26757
95.23243	103.26757
95.23243	103.26757
109.73243	117.76757
82.98243	91.01757
103.23243	111.26757
107.23243	115.26757
88.73243	96.76757
97.73243	105.76757
101.73243	109.76757
99.98243	108.01757

Les longueur des intervalles de confiance :

8.035143 8.035143 8.035143 8.035143 8.035143 8.035143 8.035143 8.035143 8.035143 8.035143
8.035143 8.035143 8.035143 8.035143 8.035143 8.035143 8.035143 8.035143 8.035143 8.035143

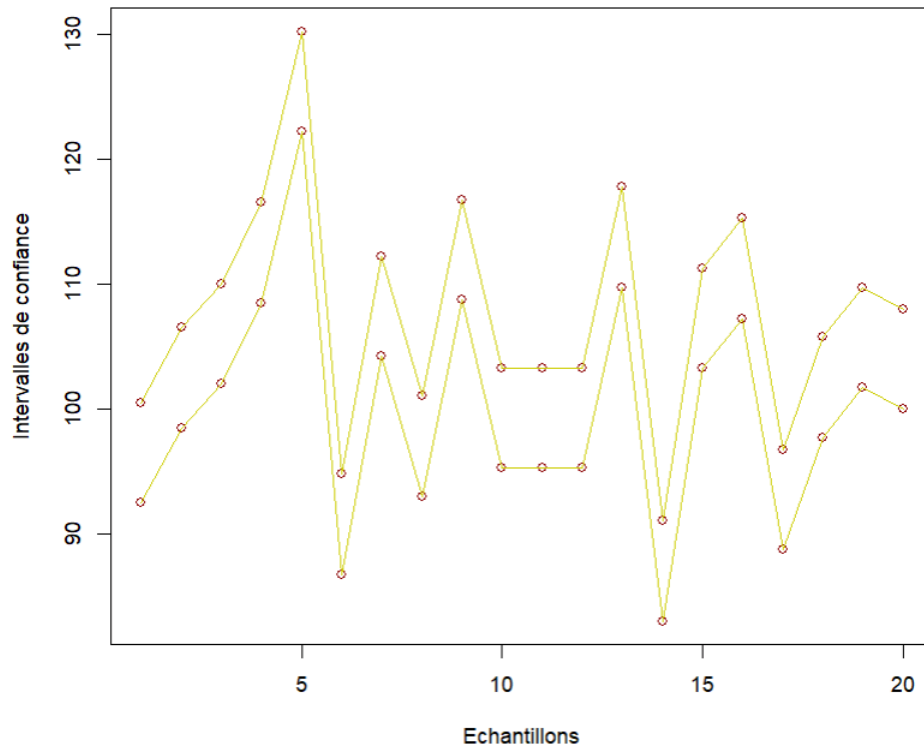


FIGURE 3.23 – **Intervalle de confiance pour stratification à allocation optimales**

La moyenne des 20 estimateurs de la moyenne vaut : **103.6875**.

3.2.4 Sondage à plusieurs degrés

Sondage par grappes

Les budgets mensuels alloués à la pratique du sport dans les 20 échantillons tirés :

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]	[,10]
[1,]	180	80	180	180	80	100	180	100	180	80
[2,]	75	90	75	75	90	100	75	100	75	90
[3,]	150	150	150	150	150	150	150	150	150	150
[4,]	250	170	250	250	170	65	250	65	250	170
[5,]	150	130	150	150	130	120	150	120	150	130
[6,]	110	150	110	110	150	90	110	90	110	150
[7,]	120	150	120	120	150	110	120	110	120	150
[8,]	40	190	40	40	190	55	40	55	40	190
[9,]	90	70	90	90	70	65	90	65	90	70
[10,]	120	40	120	120	40	65	120	65	120	40
[11,]	130	120	130	130	120	65	130	65	130	120
[12,]	140	180	140	140	180	65	140	65	140	180
[13,]	120	180	120	120	180	120	120	120	120	180
[14,]	150	70	150	150	70	90	150	90	150	70
[15,]	130	140	130	130	140	90	130	90	130	140
[16,]	140	120	140	140	120	65	140	65	140	120
[17,]	80	80	80	100	80	65	80	65	100	80
[18,]	90	90	90	100	90	60	90	60	100	90
[19,]	150	120	150	150	120	40	150	40	150	120
[20,]	170	40	170	65	40	65	170	65	65	40
[21,]	130	200	130	120	200	65	130	65	120	200
[22,]	150	100	150	90	100	65	150	65	90	100
[23,]	150	100	150	110	100	120	150	120	110	100
[24,]	190	150	190	55	150	180	190	180	55	150
[25,]	70	65	70	65	65	75	70	75	65	65
[26,]	40	120	40	65	120	150	40	150	65	120
[27,]	120	90	120	65	90	250	120	250	65	90
[28,]	180	110	180	65	110	150	180	150	65	110
[29,]	180	55	180	120	55	110	180	110	120	55
[30,]	70	65	70	90	65	120	70	120	90	65
[31,]	140	65	140	90	65	40	140	40	90	65
[32,]	120	65	120	65	65	90	120	90	65	65
[33,]	80	65	80	65	65	120	80	120	65	65

FIGURE 3.24 – **Echantillon pour sondage par grappe**

Les estimateurs de la moyenne du budget mensuel dans les 10 échantillons tirés :

125.8108 101.2500 125.8108 102.3077 101.2500 102.3077 125.8108 102.3077 102.3077 101.25

Les estimateurs de la variance dans les 10 échantillons tirés :

2204.880 1872.238 2204.880 1852.429 1872.238 1852.429 2204.880 1852.429 1852.429 1872.238

Les 10 intervalles de confiance :

122.07908	129.5425
97.82796	104.6720
122.07908	129.5425
98.67290	105.9425
97.82796	104.6720
98.67290	105.9425
122.07908	129.5425
98.67290	105.9425
98.67290	105.9425
97.82796	104.6720

Les longueur des intervalles de confiance :

7.463468 6.844082 7.463468 7.269579 6.844082 7.269579 7.463468 7.269579 7.269579 6.844082

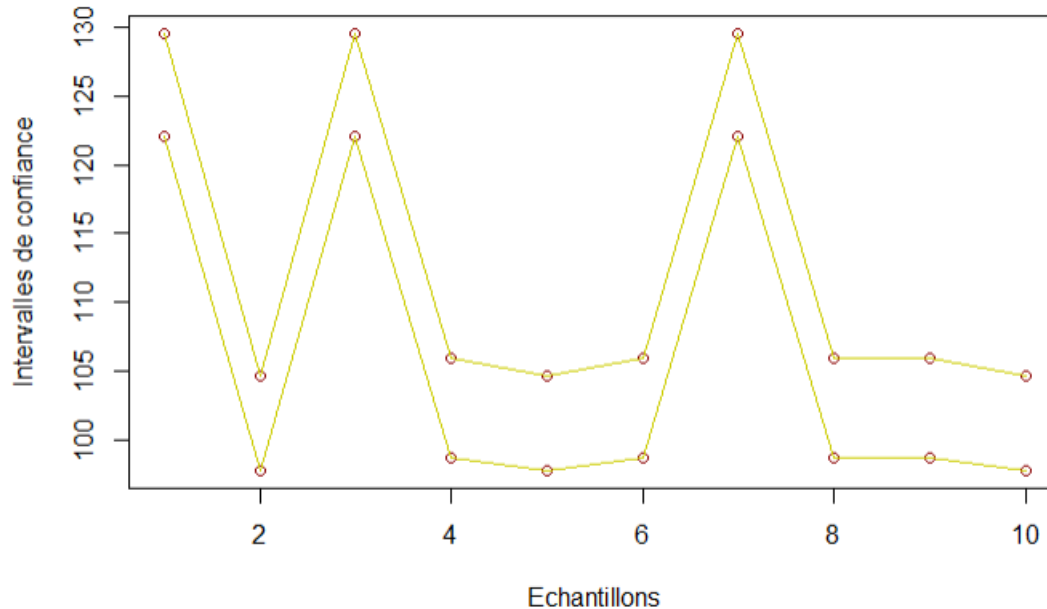


FIGURE 3.25 – **Intervalle de confiance pour sondage par grappe**

La moyenne des 10 estimateurs de la moyenne est : **109.0413**.

Sondage à deux degrés

Les budgets mensuels dans les 10 échantillons tirés :

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]	[,10]
[1,]	100	100	100	100	180	180	100	180	100	180
[2,]	100	100	100	100	75	75	100	75	100	75
[3,]	150	150	150	150	150	150	150	150	150	150
[4,]	65	65	65	65	250	250	65	250	65	250
[5,]	120	120	120	120	150	150	120	150	120	150
[6,]	90	90	90	90	110	110	90	110	90	110
[7,]	110	110	110	110	120	120	110	120	110	120
[8,]	55	55	55	55	40	40	55	40	55	40
[9,]	65	65	65	65	90	90	65	90	65	90
[10,]	65	65	65	65	120	120	65	120	65	120
[11,]	65	65	65	65	130	130	65	130	65	130
[12,]	65	65	65	65	140	140	65	140	65	140
[13,]	120	120	120	120	120	120	120	120	120	120
[14,]	90	90	90	90	150	150	90	150	90	150
[15,]	90	90	90	90	130	130	90	130	90	130
[16,]	65	65	65	65	140	140	65	140	65	140
[17,]	65	65	65	65	0	0	65	0	65	0
[18,]	60	60	60	60	0	0	60	0	60	0
[19,]	40	40	40	40	0	0	40	0	40	0
[20,]	65	65	65	65	0	0	65	0	65	0
[21,]	65	65	65	65	0	0	65	0	65	0
[22,]	65	65	65	65	0	0	65	0	65	0
[23,]	120	120	120	120	0	0	120	0	120	0

FIGURE 3.26 – **Echantillon pour sondage à deux degrés**

Les estimateurs de la moyenne :

82.3913 102.3077 82.3913 82.3913 130.9375 130.9375 82.3913 130.9375 82.3913 130.9375

Les estimateurs de la variance :

747.4308 1852.4291 747.4308 747.4308 2114.0625 2114.0625 747.4308 2114.0625 747.4308 2114.0625

Les 10 intervalles de confiance :

72.53027	92.25234
94.73493	109.88045
72.53027	92.25234
72.53027	92.25234
119.11453	142.76047
119.11453	142.76047
72.53027	92.25234
119.11453	142.76047
72.53027	92.25234
119.11453	142.76047

Les longueur des intervalles de confiance :

19.72208 15.14552 19.72208 19.72208 23.64594 23.64594 19.72208 23.64594 19.72208 23.64594

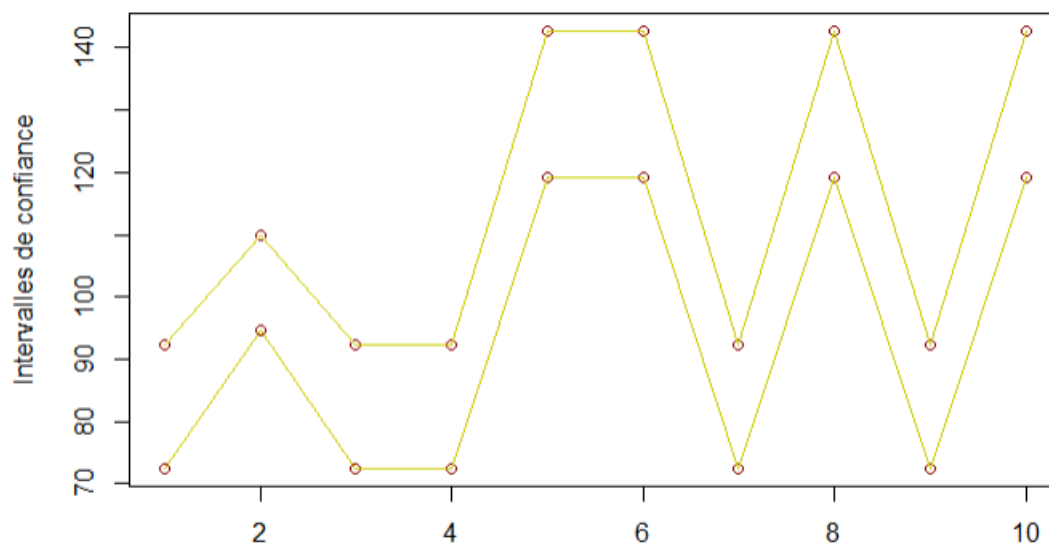


FIGURE 3.27 – Intervalle de confiance pour sondage à deux degrés

La moyenne des 10 estimateurs de la moyenne est : **103.8014**.

3.2.5 Synthèse

TABLE 3.1: **Tableau récapitulatif**

Type de Sondage	Moyenne
PESR	100.3625
PEAR	101.1875
PISR	118.525
PIAR	124.2
STRATE MEME TAILLE	104.125
STRATE PROP	101.925
STRATE OPT	103.6875
GRAPPE	109.0413
2DEGRE	103.8014

Conclusion

Pour conclure on peut dire que cette enquête nous a permis d'estimer le budget mensuel alloué à la pratique du sport dans la population Tunisienne à travers les différentes méthodes de sondage.

Chacune de ces dernières nous a fourni un tas d'informations avec un degré de précision que nous avons pu toucher à travers la demie-longueur de l'intervalle de confiance.

Nous constatons que le budget mensuel moyen dans la population est de **103.25** alors qu'il est de **103.6875** dans le cas de sondage à allocations optimales.

On peut le considérer comme le plus précis car la moyenne dans l'échantillon obtenu par cette méthode est le plus proche de la moyenne du budget mensuel de la population.

3.2.6 Redressement

Sondage de type PESR :

L'estimateur de la moyenne avant redressement par quotion est **107.25**. Or on connaît que la proportion des femmes dans la population vaut **0.51**. Alors que la proportion des femmes dans l'échantillon est **0.45**.

L'estimateur de la moyenne apres redressement par quotion est : ' **121.55**.

Annexe

```
library(ggplot2)
library(readxl)
Lecture des données
sport <- read_excel("C : /Users/syrinesbaa/Desktop/Pc
Syrine/downloads/sport2.xlsx")
View(sport)
Affichage des premeières valeurs de la base
head(sport)
attach(sport)
summary(sport) Analyse statistiques descriptives
table(sportSalle)sexe :
install.packages("plotrix")
library(plotrix)
freq=table(sportSexe)/100
x=c(freq[1],freq[2])
names(x)=c("Femme 51
pie3D(freq,labels=names(x),col=c("firebrick1","deepskyblue3")) age
freq1=table(sportAutresDPenses)/100
xp=c(1,2)
names(xp)=c("Oui (49
pie(freq1,labels=names(xp),col=c("deepskyblue3","firebrick1")) categorie socio pro
freq1=table(sportCategorie_socio_professionnelle)/100
xp=c(1,2,3)
names(xp)=c(" À la retraite/inactif(ve) (6d'emploi (50
pie(freq1,labels=names(xp),col=c("deeppink4","darksalmon","cornsilk3")) statut
freq1=table(sportStatut_Matrimonial)/100
xp=c(1,2,3)
names(xp)=c(" Célibataire (71"," Marié(e) (27
pie(freq1,labels=names(xp),col=c("darkolivegreen2","darkgoldenrod3","darkseagreen")) HEURES
freq1=table(sportHeures)/100
xp=c(1,2,3,4)
names(xp)=c(" Plus de 7 heures (16
pie(freq1,labels=names(xp),col=c("indianred1","lightpink1","ivory3","khaki")) est ce qu'ils ont des
depenses hors abonnement ou pas
```

```

freq1=table(sportAutresDPenses)/100
xp=c(1,2)
names(xp)=c("Oui (49
pie(freq1,labels=names(xp),col=c("indianred1","lightpink1")) budget
freq1=table(sport2BudgetMensuel)/100
xp=c(1,2,3)
names(xp)=c("Entre 50 et 100 Dinars (39
pie(freq1,labels=names(xp),col=c("indianred1","azure","khaki"))
library(MASS)
library(lpSolve)
library(sampling)
library(ggplot2)
library(readxl)
Lecture des données
Sondage <- read_xcel("C : /Users/Mars/Downloads/sport2 - 1.xlsx")
attach(Sondage)
p<-1 :100 pour identifier les individus de la population
Sondage<-cbind(Sondage,p)
View(Sondage)
Création de la variable d'intérêt : Budgetmensuel
y<-SondageBudgetMensuel
y
N<-100 taille de la population
n<-20
calcul de la moyenne et de la variance dans la population
moy<-mean(y)
v<-var(y)
moy le budget mensuel moyen
v la variance
SAS : PESR et PEAR Selection de l'échantillon sans remise
Mtech <- matrix(0,nrow = 20,ncol = 20)Matricechantillon(pour
stocker les 20 échantillon)
IC=matrix(0,nrow = 20,ncol = 2) Pour les 20 intervalles de confiance
ypesr <- -c()pourles20estimateursdelamoyenne
vpesr <- -c()pourles20estimateursdelavariance
Tirage de 20 échantillons sans remise
for (i in 1 :20)
s<-srswor(n,N)
Mtech[, i] <- -y[s == 1][1 : 20]
ypesr[i] <- -mean(Mtech[, i])
vpesr[i] <- -var(Mtech[, i])
Intervalles de confiance

```

```

for (i in 1 :20)
  IC[i,]=c( $y_{pesr}[i] - 1.96 * \sqrt{var(y_{pesr})/20}$ ),
 $y_{pesr}[i] + 1.96 * \sqrt{var(y_{pesr})/20}$ )
  l<-IC[,2]-IC[,1]
  moyypesr = mean( $y_{pesr}$ )/round( $Mt_{ech}$ ,1)les20chantillonssonttirsetvoicilesrsultats44
  ypesrlesestimateursdelamoyenne
  vpesrlesestimateursdelavarience
  IC les 20 intervalles de confiance
  l longueur des intervalles de confiance
  Représentation graphique des Intervalles de confiance
  plot(seq(1,20,by=1),IC[,1],col="red4",xlab="Echantillons",ylab="Intervalles de confiance", ylim =
c(min(min(IC[,2]),min(IC[,1])),max(max(IC[,2]),max(IC[,1]))))
  lines(seq(1,20,by=1),IC[,1],col="green")
  points(seq(1,20,by=1),IC[,2],col="red4")
  lines(seq(1,20,by=1),IC[,2],col="green")
  moyypesrmoymoyennedes20estimateurs
  Selection de E avec remise
  ech1<-matrix(0,nrow = 20,ncol = 20) Matrice échantillon (pour stocker les 20 échantillon)
  IC=matrix(0,nrow = 20,ncol = 2) Pour les 20 intervalles de confiance
  ypear < -c()pourles20estimateursdelamoyenne
  vpear < -c()pourles20estimateursdelavarience
  Tirage de 20 échantillons avec remise
  for (j in 1 :20)
    ech1[,j]<-sample(y,20,replace=TRUE)
    ypear[j] < -mean(ech1[,j])
    vpear[j] < -var(ech1[,j])
    les intervalles de confiances
    for (j in 1 :20)
      IC[j,]=c( $y_{pear}[j] - 1.96 * \sqrt{var(y_{pear})/20}$ ),
 $y_{pear}[j] + 1.96 * \sqrt{var(y_{pear})/20}$ )
      l1<-IC[,2]-IC[,1] calcul du longueur de l'intervalle
      moyypear = mean( $y_{pear}$ )
      round(ech1,1) les 20 échantillons sont tirés et les résultats
      ypearlesestimateursdelamoyenne
      vpearlesestimateursdelavarience
      IC les 20 intervalles de confiance
      l1 longueur des intervalles de confiance
      Représentation graphique des Intervalles de confiance
      plot(seq(1,20,by=1),IC[,1],col="red4",xlab = "Echantillons" , ylab = "Intervalles de confiance", ylim
= c(min(min(IC[,2]),min(IC[,1])),max(max(IC[,2]),max(IC[,1]))))
      lines(seq(1,20,by=1),IC[,1],col="blue4")
      points(seq(1,20,by=1),IC[,2],col="red4")

```

```

lines(seq(1,20,by=1),IC[,2],col="blue4")
moyypearmoyennedes20estimateurs
Sondage aléatoire à probabilités inégales
PISR tirage sans remise
Nous ferons le sondage en utilisant la méthode systématique
Calcul des probabilités d'inclusion
 $p_i < -inclusionprobabilities(y, 20)$ 
sum(pi)onvriefiequelasommeestgalen
déclaration
ech2<-matrix(0,nrow = 20,ncol = 20) Matrice échantillon (pour stocker les 20 échantillon)
IC=matrix(0,nrow = 20,ncol = 2) Pour les 20 intervalles de confiance
 $y_{p}^{isr} < -c()$ pourles20estimateursdelamoyenne
 $v_{p}^{isr} < -c()$ pourles20estimateursdelavariance
Tirage de 20 échantillons sans remise
for (e in 1 :20)
s=UPsystematic(pi)
ech2[,e]<-y[s==1][1 :20]
 $y_{p}^{isr}[e] < -mean(ech2[, e])$ 
 $v_{p}^{isr}[e] < -var(ech2[, e])$ 
les intervalles de confiances
for (j in 1 :20)
IC[j,]=c( $y_{p}^{isr}[j] - 1.96 * sqrt(var(y_{p}^{isr})/20)$ ,  $y_{p}^{isr}[j] + 1.96 * sqrt(var(y_{p}^{isr})/20)$ )
l2<-IC[,2]-IC[,1] calcul du longueur de l'intervalle
moyypisr < -mean(ypisr)
Resultats sont :
round(ech2,1) le budget mensuel dans les échantillons tirés
ypisrleseestimateursdelamoyenne
ypisrleseestimateursdelavariance
IC les 20 intervalles de confiance
l2 longueur des intervalles de confiance
Représentation graphique des Intervalles de confiance
plot(seq(1,20,by=1),IC[,1],col="red4",
xlab="Echantillons",ylab="Intervalles de confiance",
ylim = c(min(min(IC[,2]),min(IC[,1])),max(max(IC[,2]),max(IC[,1]))))
lines(seq(1,20,by=1),IC[,1],col="orange")
points(seq(1,20,by=1),IC[,2],col="red4")
lines(seq(1,20,by=1),IC[,2],col="orange")
moyypisrmoyennedes20estimateurs
PIAR Avec Remise
 $p_j < -inclusionprobabilities(y, 20)/n$ 
sum(pj)
v=c(0,cumsum(pj))

```

v

ech3<-matrix(0,nrow = 20,ncol = 20) Matrice échantillon (pour stocker les 20 échantillon)

IC=matrix(0,nrow = 20,ncol = 2) Pour les 20 intervalles de confiance

$y_{piar} < -c()$ pour les 20 estimateurs de la moyenne

$v_{piar} < -c()$ pour les 20 estimateurs de la variance

Tirage de 20 échantillons avec remise

for (e in 1 :20)

s=NULL

for (i in 1 :20) u<-runif(1,0,1)

for (j in 2 :101)

if((v[j-1]<=u)(v[j]>u))

s<-c(s,j-1)

ech3[,e]<-y[s]

$y_{piar}[e] < -mean(ech3[,e])$

$v_{piar}[e] < -var(ech3[,e])$

les intervalles de confiances

for (j in 1 :20)

$IC[j,]=c(y_{piar}[j] - 1.96 * sqrt(var(y_{piar})/20), y_{piar}[j] + 1.96 * sqrt(var(y_{piar})/20))$

l3<-IC[,2]-IC[,1] calcul du longueur de l'intervalle

$moy_{ypiar} < -mean(y_{piar})$

Resultats sont ci-dessous :

round(ech3,1) l'IMC dans les échantillons tirés

y_{piar} les estimateurs de la moyenne

v_{piar} les estimateurs de la variance

IC les 20 intervalles de confiance

l3 longueur des intervalles de confiance

Représentation graphique des Intervalles de confiance

plot(seq(1,20,by=1),IC[,1],col="red4",

xlab = "Echantillons",ylab = "Intervalles de confiance",

ylim = c(min(min(IC[,2]),min(IC[,1])),max(max(IC[,2]),max(IC[,1]))))

lines(seq(1,20,by=1),IC[,1],col="green")

points(seq(1,20,by=1),IC[,2],col="red4")

lines(seq(1,20,by=1),IC[,2],col="green")

moy_{ypiar} moyennes des 20 estimateurs

3.La stratification

Strates de même taille

La var sexe sera utilisée comme var de stratification

N1<-sum(Sexe=="Homme")

N2<-sum(Sexe=="Femme")

ech4<-matrix(0,nrow = 20,ncol = 20) Matrice échantillon (pour stocker les 20 échantillon)

IC=matrix(0,nrow = 20,ncol = 2) Pour les 20 intervalles de confiance

$y_{strat1} < -c()$ pour les 20 estimateurs de la moyenne

```

v_strat1 <- -c()pour les 20 estimateurs de la variance
Tirage de 20 échantillons
for (i in 1 :20)
  s<- strata(Sondage[order(SondageSexe), ], stratanames = "Sexe", size = c(10, 10), method =
"srswor")
  ech4[,i]<-y[sID_unit]
  y_strat1[i] < -mean(ech4[, i])
  v_strat1[i] < -var(ech4[, i])
Intervalles de confiance
for (j in 1 :20)
  IC[j,]=c(y_strat1[j] - 1.96 * sqrt(var(y_strat1)/20), y_strat1[j] + 1.96 * sqrt(var(y_strat1)/20))
  l4<-IC[,2]-IC[,1] calcul du longueur de l'intervalle
  moy_y_strat1 < -mean(y_strat1)
Resultats sont ci-dessous :
round(ech4,1) le budget mensuel dans les échantillons tirés
y_strat1le se estimateurs de la moyenne
v_strat1le se estimateurs de la variance
IC les 20 intervalles de confiance
l4 longueur des intervalles de confiance
Représentation graphique des Intervalles de confiance
plot(seq(1,20,by=1),IC[,1],col="red4", xlab = "Echantillons",ylab = "Intervalles de confiance", ylim
= c(min(min(IC[,2]),min(IC[,1])),max(max(IC[,2]),max(IC[,1]))))
  lines(seq(1,20,by=1),IC[,1],col="yellow")
  points(seq(1,20,by=1),IC[,2],col="red4")
  lines(seq(1,20,by=1),IC[,2],col="yellow")
  moy_y_strat1moyenne des 20 estimateurs
Stratification à allocations proportionnelles
n1<-round((n/N)*N1)
n2<-round((n/N)*N2)
ech5<-matrix(0,nrow = 20,ncol = 20) Matrice échantillon (pour stocker les 20 échantillon)
IC=matrix(0,nrow = 20,ncol = 2) Pour les 20 intervalles de confiance
y_statP < -c()pour les 20 estimateurs de la moyenne
v_statP < -c()pour les 20 estimateurs de la variance
Tirage de 20 échantillons
for (i in 1 :20)
  s<-strata(Sondage[order(SondageSexe), ], stratanames = "Sexe", size = c(n1, n2), method =
"srswor")
  ech5[,i]<-y[sID_unit]
  y_statP[i] < -mean(ech5[, i])
  v_statP[i] < -var(ech5[, i])
Intervalles de confiance
for (j in 1 :20)

```

```

IC[j,]=c(ysstatP[j] - 1.96 * sqrt(var(ysstatP)/20), ysstatP[j] + 1.96 * sqrt(var(ysstatP)/20))
l5<-IC[,2]-IC[,1] calcul du longeur de l'intervalle
moyysstatP < -mean(ysstatP)
Resultats sont ci-dessous :
round(ech5,1) le budget mensuel dans les échantillons tirés
ysstatPlesestimateursdelamoyenne
vsstatPlesestimateursdelavarariance
IC les 20 intervalles de confiance
l5 longueur des intervalles de confiance
Représentation graphique des Intervalles de confiance
plot(seq(1,20,by=1),IC[,1],col="red4", xlab = "Echantillons",ylab = "Intervalles de confiance", ylim
= c(min(min(IC[,2]),min(IC[,1])),max(max(IC[,2]),max(IC[,1]))))
lines(seq(1,20,by=1),IC[,1],col="yellow4")
points(seq(1,20,by=1),IC[,2],col="red4")
lines(seq(1,20,by=1),IC[,2],col="yellow4")
moyysstatP
Allocation Optimale
yf < -SondageBudgetMensuel
yh < -SondageBudgetMensuel
Taille des strates
nf<-round((20*N1*sqrt((n/(n-1))*var(yf)))/(N1*sqrt((n/(n-1))*var(yf))+N2*sqrt((n/(n-
1)) * var(yh))))nf
nh<-round((20*N1*sqrt((n/(n-1))*var(yh)))/(N1*sqrt((n/(n-1))*var(yf))+N2*sqrt((n/(n-
1)) * var(yh))))nh
déclaration
ech6<-matrix(0,nrow = 20,ncol = 20) Matrice échantillon (pour stocker les 20 échantillon)
IC=matrix(0,nrow = 20,ncol = 2) Pour les 20 intervalles de confiance
ystratopt < -c()pourles20estimateursdelamoyenne
vstratopt < -c()pourles20estimateursdelavarariance
Tirage de 20 échantillons
for (i in 1 :20)
s<-strata(Sondage[order(Sondage.Sexe),], stratanames = "Sexe", size = c(nf, nh), method =
"srswor")ech6[,i] < -y[sIDunit]
ystratopt[i] < -mean(ech6[,i])
vstratopt[i] < -var(ech6[,i])
Intervalles de confiance
for (j in 1 :20)
IC[j,]=c(ystratopt[j]-1.96*sqrt(var(ystratopt)/20), ystratopt[j]+1.96*sqrt(var(ystratopt)/20))
l6<-IC[,2]-IC[,1] calcul du longeur de l'intervalle
moyystratopt < -mean(ystratopt)
Resultats sont ci-dessous :
round(ech6,1) le budget mensuel dans les échantillons tirés

```

```

ystratoptlesestimeursdelamoyenne
vstratoptlesestimeursdelavariance
IC les 20 intervalles de confiance
l6 longueur des intervalles de confiance
Représentation graphique des Intervalles de confiance
plot(seq(1,20,by=1),IC[,1],col="red4",
xlab = "Echantillons",ylab = "Intervalles de confiance",
ylim = c(min(min(IC[,2]),min(IC[,1])),max(max(IC[,2]),max(IC[,1]))))
lines(seq(1,20,by=1),IC[,1],col="yellow3")
points(seq(1,20,by=1),IC[,2],col="red4")
lines(seq(1,20,by=1),IC[,2],col="yellow3")
moyystratopt
Sondage à plusieurs degré
SONDAGE PAR GRAPPE
ech7<-matrix(0,nrow =60,ncol=10) Matrice qui contiendra les 10 échantillons
IC<-matrix(0,nrow = 10,ncol=2) Matrice pour les 10 intervalles de confiances
ygrap < -c()pourstockerles10estimateursdelamoyenne
vgrap < -c()pourlavariance
Nous allons travailler avec la valeur auxiliaire "Heures" temps consacrés au sport
m<-c("Entre 1 et 3 heures","Entre 3 et 5 heures ","Entre 5 et 7 heures","Plus de 7 heures")
M=4
1er tirage
ech71<-c() le vecteur du premier tirage
s=sample(m,2)
ech71<-c(y[Heures==s[1]],y[Heures==s[2]])
ech7[1:length(ech71),1]<-ech71 matrice ech7
répétition de l'opération 10 fois
ygrap[1] < -mean(ech71)
vgrap[1] = var(ech71)
2è tirage
ech72<-c()
s<-sample(m,2)
ech72<-c(y[Heures==s[1]],y[Heures==s[2]])
ech7[1:length(ech72),2]<-ech72
ygrap < -c(ygrap, mean(ech72))
vgrap < -c(vgrap, var(ech72))
3ème tirage
ech73<-c()
s<-sample(m,2)
ech73<-c(y[Heures==s[1]],y[Heures==s[2]])
ech7[1:length(ech73),3]=ech73
ygrap < -c(ygrap, mean(ech73))

```



```
vgrap < -c(vgrap, var(ech73))
4e tirage
ech74<-c()
s<-sample(m,2)
ech74<-c(y[Heures==s[1]],y[Heures==s[2]])
ech7[1 :length(ech74),4]=ech74
ygrap < -c(ygrap, mean(ech74))
vgrap < -c(vgrap, var(ech74))
5e tirage
ech75<-c()
s<-sample(m,2)
ech75<-c(y[Heures==s[1]],y[Heures==s[2]])
ech7[1 :length(ech75),5]=ech75
ygrap < -c(ygrap, mean(ech75))
vgrap < -c(vgrap, var(ech75))
6e tirage
ech76<-c()
s<-sample(m,2)
ech76<-c(y[Heures==s[1]],y[Heures==s[2]])
ech7[1 :length(ech76),6]=ech76
ygrap < -c(ygrap, mean(ech76))
vgrap < -c(vgrap, var(ech76))
7e tirage
ech77<-c()
s<-sample(m,2)
ech77<-c(y[Heures==s[1]],y[Heures==s[2]])
ech7[1 :length(ech77),7]=ech77
ygrap < -c(ygrap, mean(ech77))
vgrap < -c(vgrap, var(ech77))
8e tirage
ech78<-c()
s<-sample(m,2)
ech78<-c(y[Heures==s[1]],y[Heures==s[2]])
ech7[1 :length(ech78),8]=ech78
ygrap < -c(ygrap, mean(ech78))
vgrap < -c(vgrap, var(ech78))
9e tirage
ech79<-c()
s<-sample(m,2)
ech79<-c(y[Heures==s[1]],y[Heures==s[2]])
ech7[1 :length(ech79),9]=ech79
ygrap < -c(ygrap, mean(ech79))
```

```

vgrap <- -c(vgrap, var(ech79))
10e tirage
ech75<-c()
s<-sample(m,2)
ech710<-c(y[Heures==s[1]],y[Heures==s[2]])
ech7[1:length(ech710),10]=ech710
ygrap <- -c(ygrap, mean(ech710))
vgrap <- -c(vgrap, var(ech710))
Intervalles de confiance
for (j in 1:10)
IC[j,]=c(ygrap[j]-1.96*sqrt(var(ygrap)/sum(ech7[,j]!=0)), ygrap[j]+1.96*sqrt(var(ygrap)/sum(ech7[,j]!=0)))
l7<-IC[,2]-IC[,1] calcul du longeur de l'intervalle
moyygrap <- -mean(ygrap)
Resultats sont ci-dessous :
round(ech7,1) l'IMC dans les échantillons tirés
ygrapleestimateursdelamoyenne
vgrapleestimateursdelavariance
IC les 20 intervalles de confiance
l7 longueur des intervalles de confiance
Représentation graphique des Intervalles de confiance
plot(seq(1,10,by=1),IC[,1],col="red4",xlab = "Echantillons", ylab = "Intervalles de confiance", ylim
= c(min(min(IC[,2]),min(IC[,1])),max(max(IC[,2]),max(IC[,1]))))
lines(seq(1,10,by=1),IC[,1],col="yellow3")
points(seq(1,10,by=1),IC[,2],col="red4")
lines(seq(1,10,by=1),IC[,2],col="yellow3")
moyygrap
Sondage à deux degré
ech8<-matrix(0,nrow = 60,ncol=10) Matrice qui contiendra les 10 échantillons
IC<-matrix(0,nrow = 10,ncol=2) Matrice pour les 10 intervalles de confiances
y2deg <- -c()pourstockerles10estimateursdelamoyenne
v2deg <- -c()pourdelavarianceNousallonstravailleravec la variableauxiliaire" Heures"
temps consacrés au sport
cbind(Sondage,p)
1er tirage
s1<-c()
s1<-sample(m,2)
grp1<-Sondage[Heures==s1[1],]
grp2<-Sondage[Heures==s1[2],]
taille de l'échantillon tiré dans les grappes
if(round(length(grp1p)/5) == 0)nh1 = 1
if(round(length(grp1p)/5) > 0)nh1 = round(length(grp1p)/5)

```

```
if(round(length(grp1p)/5) == 0)nh2 = 1
if(round(length(grp1p)/5) > 0)nh2 = round(length(grp1p)/5)
tirage dans les grappes
s2=c(sample(grp1p, nh1), sample(grp2p,nh2))
ech81<-c()
ech81[1]=y[p==s2[1],]
for (i in (2 :length(s2)))
ech81<-c(ech81,y[p==s2[i],])
ech8[1 :length(ech81),1]=ech81
y2deg < -mean(ech81)
v2deg < -mean(ech81)
2è tirage
s1<-c()
s1<-sample(m,2)
grp1<-Sondage[Heures==s1[1],]
grp2<-Sondage[Heures==s1[2],]
taille de l'échantillon tiré dans les grappes
if(round(length(grp1p)/5) == 0)nh1 = 1
if(round(length(grp1p)/5) > 0)nh1 = round(length(grp1p)/5)
if(round(length(grp1p)/5) == 0)nh2 = 1
if(round(length(grp1p)/5) > 0)nh2 = round(length(grp1p)/5)
tirage dans les grappes
s2=c(sample(grp1p, nh1), sample(grp2p,nh2))
ech82<-c()
ech82[1]=y[p==s2[1]]
for (i in (2 :length(s2)))
ech82<-c(ech82,y[p==s2[i]])
ech8[1 :length(ech82),2]=ech82
y2deg < -c(y2deg, mean(ech82))
v2deg < -c(v2deg, var(ech82))
3è tirage
s1<-c()
s1<-sample(m,2)
grp1<-Sondage[Heures==s1[1],]
grp2<-Sondage[Heures==s1[2],]
taille de l'échantillon tiré dans les grappes
if(round(length(grp1p)/5) == 0)nh1 = 1
if(round(length(grp1p)/5) > 0)nh1 = round(length(grp1p)/5)
if(round(length(grp1p)/5) == 0)nh2 = 1
if(round(length(grp1p)/5) > 0)nh2 = round(length(grp1p)/5)
tirage dans les grappes
s2=c(sample(grp1p, nh1), sample(grp2p,nh2))
```

```
ech83<-c()
ech83[1]=y[p==s2[1]]
for (i in (2 :length(s2)))
ech83<-c(ech83,y[p==s2[i]])
ech8[1 :length(ech83),3]=ech83
 $y_2deg < -c(y_2deg, mean(ech83))$ 
 $v_2deg < -c(v_2deg, var(ech83))$ 
4è tirage
s1<-c()
s1<-sample(m,2)
grp1<-Sondage[Heures==s1[1],]
grp2<-Sondage[Heures==s1[2],]
taille de l'échantillon tiré dans les grappes
if(round(length(grp1p)/5) == 0)nh1 = 1
if(round(length(grp1p)/5) > 0)nh1 = round(length(grp1p)/5)
if(round(length(grp1p)/5) == 0)nh2 = 1
if(round(length(grp1p)/5) > 0)nh2 = round(length(grp1p)/5) tirage dans les grappes
s2=c(sample(grp1p, nh1), sample(grp2p,nh2))
ech84<-c()
ech84[1]=y[p==s2[1]]
for (i in (2 :length(s2)))
ech84<-c(ech84,y[p==s2[i]])
ech8[1 :length(ech84),4]=ech84
 $y_2deg < -c(y_2deg, mean(ech84))$ 
 $v_2deg < -c(v_2deg, var(ech84))$ 5tirage
s1<-c()
s1<-sample(m,2)
grp1<-Sondage[Heures==s1[1],]
grp2<-Sondage[Heures==s1[2],]
taille de l'échantillon tiré dans les grappes
if(round(length(grp1p)/5) == 0)nh1 = 1
if(round(length(grp1p)/5) > 0)nh1 = round(length(grp1p)/5)
if(round(length(grp1p)/5) == 0)nh2 = 1
if(round(length(grp1p)/5) > 0)nh2 = round(length(grp1p)/5)
tirage dans les grappes
s2=c(sample(grp1p, nh1), sample(grp2p,nh2))
ech85<-c()
ech85[1]=y[p==s2[1]]
for (i in (2 :length(s2)))
ech85<-c(ech85,y[p==s2[i]])
ech8[1 :length(ech85),5]=ech85
 $y_2deg < -c(y_2deg, mean(ech85))$ 
```

```
v2deg < -c(v2deg, var(ech85))
6e tirage
s1<-c()
s1<-sample(m,2)
grp1<-Sondage[Heures==s1[1],]
grp2<-Sondage[Heures==s1[2],]
taille de l'échantillon tiré dans les grappes
if(round(length(grp1p)/5) == 0)nh1 = 1
if(round(length(grp1p)/5) > 0)nh1 = round(length(grp1p)/5)
if(round(length(grp1p)/5) == 0)nh2 = 1
if(round(length(grp1p)/5) > 0)nh2 = round(length(grp1p)/5)
tirage dans les grappes
s2=c(sample(grp1p, nh1), sample(grp2p,nh2))
ech86<-c()
ech86[1]=y[p==s2[1]]
for (i in (2 :length(s2)))
ech86<-c(ech86,y[p==s2[i]])
ech8[1:length(ech86),6]=ech86
y2deg < -c(y2deg, mean(ech86))
v2deg < -c(v2deg, var(ech86))
7e tirage
s1<-c()
s1<-sample(m,2)
grp1<-Sondage[Heures==s1[1],]
grp2<-Sondage[Heures==s1[2],]
taille de l'échantillon tiré dans les grappes
if(round(length(grp1p)/5) == 0)nh1 = 1
if(round(length(grp1p)/5) > 0)nh1 = round(length(grp1p)/5)
if(round(length(grp1p)/5) == 0)nh2 = 1
if(round(length(grp1p)/5) > 0)nh2 = round(length(grp1p)/5)
tirage dans les grappes
s2=c(sample(grp1p, nh1), sample(grp2p,nh2))
ech87<-c()
ech87[1]=y[p==s2[1]]
for (i in (2 :length(s2)))
ech87<-c(ech87,y[p==s2[i]])
ech8[1:length(ech87),7]=ech87
y2deg < -c(y2deg, mean(ech87))
v2deg < -c(v2deg, var(ech87))
8e tirage
s1<-c()
s1<-sample(m,2)
```

```
grp1<-Sondage[Heures==s1[1],]
grp2<-Sondage[Heures==s1[2],]
taille de l'échantillon tiré dans les grappes
if(round(length(grp1p)/5) == 0)nh1 = 1
if(round(length(grp1p)/5) > 0)nh1 = round(length(grp1p)/5)
if(round(length(grp1p)/5) == 0)nh2 = 1
if(round(length(grp1p)/5) > 0)nh2 = round(length(grp1p)/5)
tirage dans les grappes
s2=c(sample(grp1p, nh1), sample(grp2p,nh2))
ech88<-c()
ech88[1]=y[p==s2[1]]
for (i in (2 :length(s2)))
ech88<-c(ech88,y[p==s2[i]])
ech8[1:length(ech88),8]=ech88
y2deg < -c(y2deg, mean(ech88))
v2deg < -c(v2deg, var(ech88))
9è tirage
s1<-c()
s1<-sample(m,2)
grp1<-Sondage[Heures==s1[1],]
grp2<-Sondage[Heures==s1[2],]
taille de l'échantillon tiré dans les grappes
if(round(length(grp1p)/5) == 0)nh1 = 1
if(round(length(grp1p)/5) > 0)nh1 = round(length(grp1p)/5)
if(round(length(grp1p)/5) == 0)nh2 = 1
if(round(length(grp1p)/5) > 0)nh2 = round(length(grp1p)/5)
tirage dans les grappes
s2=c(sample(grp1p, nh1), sample(grp2p,nh2))
ech89<-c()
ech89[1]=y[p==s2[1]]
for (i in (2 :length(s2)))
ech89<-c(ech89,y[p==s2[i]])
ech8[1:length(ech89),9]=ech89
y2deg < -c(y2deg, mean(ech89))
v2deg < -c(v2deg, var(ech89))
10è tirage
s1<-c()
s1<-sample(m,2)
grp1<-Sondage[Heures==s1[1],]
grp2<-Sondage[Heures==s1[2],]
taille de l'échantillon tiré dans les grappes
if(round(length(grp1p)/5) == 0)nh1 = 1
```

```

if(round(length(grp1p)/5) > 0)nh1 = round(length(grp1p)/5)
if(round(length(grp1p)/5) == 0)nh2 = 1
if(round(length(grp1p)/5) > 0)nh2 = round(length(grp1p)/5)
tirage dans les grappes
s2=c(sample(grp1p, nh1), sample(grp2p,nh2))
ech810<-c()
ech810[1]=y[p==s2[1]]
for (i in (2 :length(s2)))
ech810<-c(ech810,y[p==s2[i]])
ech8[1 :length(ech810),10]=ech810
y2deg < -c(y2deg, mean(ech810))
v2deg < -c(v2deg, var(ech810))
Intervalles de confiance
for (j in 1 :10)
IC[j,]=c(y2deg[j]-1.96*sqrt(var(y2deg)/sum(ech8[,j] != 0)), y2deg[j]+1.96*sqrt(var(y2deg)/sum(ech8
0)))
l8<-IC[,2]-IC[,1] calcul du longueur de l'intervalle
moyy2deg < -mean(y2deg)
Resultats sont ci-dessous :
round(ech8,1) l'IMC dans les échantillons tirés
y2deglesestimateursdelamoyenne
v2deglesestimateursdelavariance
IC les 20 intervalles de confiance
l8 longueur des intervalles de confiance
Représentation graphique des Intervalles de confiance
plot(seq(1,10,by=1),IC[,1],col="red4", xlab = "Echantillons",ylab = "Intervalles de confiance", ylim
= c(min(min(IC[,2]),min(IC[,1])),max(max(IC[,2]),max(IC[,1]))))
lines(seq(1,10,by=1),IC[,1],col="yellow2")
points(seq(1,10,by=1),IC[,2],col="red4")
lines(seq(1,10,by=1),IC[,2],col="yellow2")
moyy2deg
6/ Redressement par le quotient :
Sondage de type PESR
s=srswor(n,N)
echantillon=y[s==1]
moyenne=mean(echantillon)
Alors l'estimateur de la moyenne avant redressement par quotion est :
moyenne
[1] 107.25
Or on connait que la proportion des femmes dans la population est
P=sum(Sexe=="Femme")/N
P

```

[1] 0.51

Alors que la proportion des femmes dans l'échantillon est

$p = \text{sum}(\text{Sexe}[s==1] == \text{"Femme"})/n$

p

[1] 0.45

L'estimateur de la moyenne après redressement par quotient est : ‘

$\text{moyenne.red} = (P/p) * \text{moyenne}$

moyenne.red

[1] 121.55