NYA DEVAS LANELLE

SAE STATISTIQUE INFERENTIELLE

BUT
SCIENCE DES DONNEES (SD)

ANNÉE UNIVERSITAIRE 2024-2025

STATISTIQUE INFERENTIELLE

SAE 202 : Estimation par sondage simple

Introduction

Dans le cadre de notre étude statistique intitulée **"Échantillonnage par sondage simple"**, l'objectif principal est d'analyser statistiquement un ensemble de données issues d'une population donnée afin d'en tirer des estimations fiables. Plus précisément, nous visons à réaliser une **estimation par intervalle de confiance** des paramètres d'une population à partir d'un **échantillon aléatoire simple**.

Cette étude repose sur l'analyse de trois variables principales : **l'âge**, **le sexe**, et **l'avis** des individus sondés. Nous étudierons d'abord l'ensemble de la population, puis nous comparerons les résultats obtenus entre les sous-populations des **hommes** et des **femmes**.

À travers ce travail, nous appliquerons des notions clés de la statistique descriptive et inférentielle telles que l'**espérance mathématique**, la **variance** et l'**écart-type**.

PARTIE 1 :

Calcule des différents paramètres de l'âge sur la population total, population des hommes, populations des femmes.

POPULATION TOTALE	
Paramètres	Valeurs
Espérance	39,5
Variance	36,8
Écart-Type	6,1

POPULATION FEMME	
Paramètres	Valeurs
Espérance	39,7
Variance	35,9
Écart-Type	5,98

POPULATION HOMME	
Paramètres	Valeurs
Espérance	39,30
Variance	37,8
Écart-Type	6,14

PARTIE2 : Estimation de l'âge moyen de la population totale

Conception de l'échantillons (50 individus)

Soit la table des nombres aléatoires. Prenons comme point d'entrée **la première ligne, première colonne**, et comme itinéraire de lecture **du haut vers le bas**. Étant donné que la population est composée de **500 individus**, nous allons considérer **les trois derniers chiffres** de chaque nombre.

NB : Sachant que la première colonne n'a pas été suffisante pour avoir les 50 numéros j'ai utilisé la première colonne.

Les numéros sélectionnés pour notre échantillon sont :

373, 476, 269, 44, 448, 290, 106, 420, 262, 175,

511, 73, 133, 140, 38, 297, 430, 260, 231, 291,

210, 337, 20, 256, 475, 350, 179, 82, 106,

, 52, 226, 476,499, 356,297, 435, 339, 106, 430, 260

Les âges correspondants :

38, 34, 48, 38, 42, 31, 48, 34, 50, 50, 42, 36, 47, 32, 36, 37, 41, 50, 50, 40, 43, 46, 48, 46, 34, 42, 41, 31, 40, 34, 34, 32, 49, 45, 30, 32, 39, 47, 36, 41, 48, 50, 50, 40, 48, 41, 50, 48, 41, 50

Réalisation de l'estimation ponctuelle : avec la moyenne comme estimateur

On a donc pour moyenne estimée m :

$\bar{X} = 1/50 * \sum xi = 2080 / 50 = 41,6$

Estimation de m par un intervalle de confiance de 1 - alpha

Soit I l'intervalle de confiance pour m au niveau de confiance 1 - α avec α ∈ {10%, 5%, 1%}

L'intervalle est sous la forme :

$I = [\bar{X} - k ; \bar{X} + k]$

où k est la marge d'erreur.

- **Détermination de I :**

Nous devons d'abord trouver la valeur de k telle que :

$P(\bar{X} - k < m < \bar{X} + k) = 1 - \alpha$

- **Élément de contexte**

X : l'âge des individus

- **Caractérisation de la loi de la variable X et ses paramètres :**

inconue

- **Tableau :**

	Population : N = 500	Échantillon : n = 50
Espérance	$E(X) = 39{,}5$	$\bar{X} = 41{,}6$
Variance		$s2 \approx 40{,}57$

- **Conditions expérimentales :**

- Population inconnu

- Variance inconnue

- $n > 30$

Construction de I pour un niveau de confiance de 0.05, soit 95 % :

Détermination de k :

$P(\bar{X} - k < m < \bar{X} + k) = 0{,}95$

$\Rightarrow P(-k < m - \bar{X} < k) = 0{,}95$

$\Rightarrow P(-k / (\sigma/\sqrt{n}) < (\bar{X} - m) / (\sigma/\sqrt{n}) < k / (\sigma/\sqrt{n})) = 0{,}95$

Soit $T = (\bar{X} - m) / (\sigma/\sqrt{n}) \sim$ app $N(0,1)$ estimé par $s/\sqrt{n}$ $T = (\bar{X} - m) / (s/\sqrt{n})$

$\Rightarrow P(-a < T < a) = 0{,}95$ avec $a = k / (s/\sqrt{n})$

$\Rightarrow F(a) - F(-a) = 0{,}95$

$\Rightarrow 2F(a) = 1{,}95 \Rightarrow F(a) = 0{,}975 \Rightarrow$ car $F(a)=$ **$1 - \alpha/2$**

$a = 1{,}96$ (table de loi normale)

- **Calcul de k :**

$k = 1{,}96 \times (s / \sqrt{n}) = 1{,}96 \times (6{,}065 / \sqrt{50}) \approx 1{,}68$

- **Intervalle de confiance à 95 % :**

$I = [41{,}6 - 1{,}68 ; 41{,}6 + 1{,}68] = [39{,}92 ; 43{,}28]$

Conclusion :

Cela signifie que l'intervalle de confiance à 95% pour la moyenne des âges de la population se situe entre 39.92 et 43,28.

Construction de I pour un niveau de confiance de 0.1, soit 90 % :

$P(\bar{X} - k < m < \bar{X} + k) = 0,9$

$\Rightarrow P(-k < m - \bar{X} < k) = 0,9$

$\Rightarrow P(-k / (\sigma/\sqrt{n}) < (\bar{X} - m) / (\sigma/\sqrt{n}) < k / (\sigma/\sqrt{n})) = 0,9$

Soit $T = (\bar{X} - m) / (\sigma/\sqrt{n}) \sim$ app $N(0,1)$ estimé par $s/\sqrt{n}$ $T = (\bar{X} - m) / (s/\sqrt{n})$

$\Rightarrow P(-a < T < a) = 0,9$ avec $a = k / (s/\sqrt{n})$

$\Rightarrow F(a) - F(-a) = 0,9$

$\Rightarrow 2F(a) = 1,9 \Rightarrow F(a) = 0,95 \Rightarrow a = 1,645$ (table de loi normale)

- **Calcul de k :**

$k = 1,645 \times (s / \sqrt{n}) = 1,645 \times (6,065 / \sqrt{50}) \approx 1,411$

- **Intervalle de confiance à 90 % :**

$I = [41,6-1,411, 41,6+1,411]=[40,189, 43,011]$

Conclusion :

Cela signifie que l'intervalle de confiance à 90% pour la moyenne des âges de la population se situe entre 40,189 et 43,011.

Construction de I pour un niveau de confiance de 0.01, soit 99% :

$P(\bar{X} - k < m < \bar{X} + k) = 0,99$

$\Rightarrow P(-k < m - \bar{X} < k) = 0,99$

$\Rightarrow P(-k / (\sigma/\sqrt{n}) < (\bar{X} - m) / (\sigma/\sqrt{n}) < k / (\sigma/\sqrt{n})) = 0,99$

Soit $T = (\bar{X} - m) / (\sigma/\sqrt{n}) \sim$ app $N(0,1)$ estimé par $s/\sqrt{n}$ $T = (\bar{X} - m) / (s/\sqrt{n})$

$\Rightarrow P(-a < T < a) = 0,99$ avec $a = k / (s/\sqrt{n})$

$\Rightarrow F(a) - F(-a) = 0,99$

$\Rightarrow 2F(a) = 1,99$

$F(a)= 0,995$

$a=2,576$

- **Calcul de k :**

$\sigma = \sqrt{36,8} \approx 6,065$

$k = 2,576 \times (s / \sqrt{n}) = 2,576 \times (6,065 / \sqrt{50}) \approx 2,213$

- **Intervalle de confiance à 90 % :**

$I = [41,6-2,213, 41,6+2,213] = [39,387, 43,813]$

Conclusion :

Cela signifie que l'intervalle de confiance à 99% pour la moyenne des âges de la population se situe entre 39,387et 43,813.

Estimation de l’âge moyen de la population des femmes :

Conception de l’échantillon (30 individus)

Soit la table des nombres aléatoires. Prenons comme point d'entrée **la première ligne, 3 3éme colonne**, et comme itinéraire de lecture **du haut vers le bas**.

NB : 4-ème colonne utilise si on pas assez de valeurs ainsi de suite
Étant donné que la population est composée de **500 individus**, nous allons considérer **les trois derniers chiffres** de chaque nombre.

Il faut déjà remarquer qu'on a les femmes pour les numéros de ligne suivant, donc mon choix de numéros aléatoires doit respecter ces numéros de ligne :142, 8, 380, 12, 34, 161, 36, 145, 484, 396,218,389,291,157,191,255,135,426,122,50,116, 200, 205, 283, 482, 48, 76, 384, 53, 183

Les âges correspondants : 50, 48, 43, 33, 30, 36, 40, 39, 45, 42, 33, 50, 40, 30, 49, 49, 30, 46, 31, 34, 42, 47, 49, 36, 39, 50, 47, 31, 39, 40

Réalisation de l'estimation ponctuelle : avec la moyenne comme estimateur

On a donc pour moyenne estimée m :

$\bar{X} = 1/50 * \sum xi = 1268 / 30 = 42.267$

Estimation de m par un intervalle de confiance de 1 - alpha

Soit I l'intervalle de confiance pour m au niveau de confiance 1 - α avec α ∈ {10%, 5%, 1%}

L'intervalle est sous la forme :

$I = [\bar{X} - k ; \bar{X} + k]$

où k est la marge d'erreur.

- **Détermination de I :**

Nous devons d'abord trouver la valeur de k telle que :

$\mathbf{P(\bar{X} - k < m < \bar{X} + k) = 1 - \alpha}$

- **Élément de contexte**

X : l'âge des individus

- **Caractérisation de la loi de la variable X et ses paramètres :**

inconnue

- **Tableau :**

	Population : N = 245	Échantillon : n = 30
Espérance	E(X) = 39,7	$\bar{X}$ = 42,267

Variance		s2≈35,9

- **Conditions expérimentales :**

- Population inconnue

- Variance inconnue

- $n > 30$

Construction de I pour un niveau de confiance de 0.05, soit 95 % :

Détermination de k :

$P(\bar{X} - k < m < \bar{X} + k) = 0,95$

$\Rightarrow P(-k < m - \bar{X} < k) = 0,95$

$\Rightarrow P(-k / (\sigma/\sqrt{n}) < (\bar{X} - m) / (\sigma/\sqrt{n}) < k / (\sigma/\sqrt{n})) = 0,95$

Soit $T = (\bar{X} - m) / (\sigma/\sqrt{n}) \sim$ app $N(0,1)$ estimé par $s/\sqrt{n}$ on a $T = (\bar{X} - m) / (s/\sqrt{n})$

$\Rightarrow P(-a < T < a) = 0,95$ avec $a = k / (s/\sqrt{n})$

$\Rightarrow F(a) - F(-a) = 0,95$

$\Rightarrow 2F(a) = 1,95 \Rightarrow F(a) = 0,975 \Rightarrow$ car $F(a)=$ **$1 - \alpha/2$**

$a = 1,96$ (table de loi normale)

- **Calcul de k :**

$k = 1,96 \times (s / \sqrt{n}) = 1,96 \times (5,98 / \sqrt{30}) \approx 2.14$

- **Intervalle de confiance à 95 % :**

$I = [42.27 – 2.14 ; 42.27 + 2.14] = [40.13 ; 44.41]$

Conclusion :

Cela signifie que l'intervalle de confiance à 95% pour la moyenne des âges de la population se situe entre 40.13 et 44.41.

Construction de I pour un niveau de confiance de 0.1, soit 90 % :

$P(\bar{X} - k < m < \bar{X} + k) = 0,9$

$\Rightarrow P(-k < m - \bar{X} < k) = 0,9$

$\Rightarrow P(-k / (\sigma/\sqrt{n}) < (\bar{X} - m) / (\sigma/\sqrt{n}) < k / (\sigma/\sqrt{n})) = 0,9$

Soit $T = (\bar{X} - m) / (\sigma/\sqrt{n}) \sim$ app $N(0,1)$ estimé par $s/\sqrt{n}$ on a $T = (\bar{X} - m) / (s/\sqrt{n})$

$\Rightarrow P(-a < T < a) = 0,9$ avec $a = k / (s/\sqrt{n})$

$\Rightarrow F(a) - F(-a) = 1,9$

$\Rightarrow 2F(a) = 1,9 \Rightarrow F(a) = 0,95 \Rightarrow a = 1,645$ (table de loi normale)

- **Calcul de k :**

$\sigma = \sqrt{36,8} \approx 6,065$

$k = 1,645 \times (s / \sqrt{n}) = 1,645 \times (5.98 / \sqrt{30}) \approx 1,796$

- **Intervalle de confiance à 90 % :**

$I = [42,27-1,796, 42,27+1,796]=[40,474, 44,066]$

Conclusion :

Cela signifie que l'intervalle de confiance à 90% pour la moyenne des âges de la population se situe entre 40,47 et 44,066.

Construction de I pour un niveau de confiance de 0.01, soit 99% :

$P(\bar{X} - k < m < \bar{X} + k) = 0,99$

$\Rightarrow P(-k < m - \bar{X} < k) = 0,99$

$\Rightarrow P(-k / (\sigma/\sqrt{n}) < (\bar{X} - m) / (\sigma/\sqrt{n}) < k / (\sigma/\sqrt{n})) = 0,99$

Soit $T = (\bar{X} - m) / (\sigma/\sqrt{n}) \sim$ app $N(0,1)$ estimé par $s/\sqrt{n}$ on a $T = (\bar{X} - m) / (s/\sqrt{n})$

$\Rightarrow P(-a < T < a) = 0,99$ avec $a = k / (s/\sqrt{n})$

$\Rightarrow F(a) - F(-a) = 0,99$

$\Rightarrow 2F(a) = 0,99$

$F(a)= 0,995$

$a=2,576$

- **Calcul de k :**

$k = 2,576 \times (s / \sqrt{n}) = 2,576 \times (5.98/ \sqrt{30}) \approx 2.812$

- **Intervalle de confiance à 90 % :**

$I = [42,27-2,812, 42.27+2,812]=[39,458, 45.082]$

Conclusion :

Cela signifie que l'intervalle de confiance à 99% pour la moyenne des âges de la population se situe entre 39,458 et 45,082.

Estimation de l'âge moyen de la population des hommes :

Conception de l'échantillon (30 individus)

Soit la table des nombres aléatoires. Prenons comme point d'entrée **la première ligne, 3éme colonne**, et comme itinéraire de lecture **du haut vers le bas**. Étant donné que la population est composée de **500 individus**, nous allons considérer **les trois derniers chiffres** de chaque nombre.

Il faut déjà remarquer qu'on a les hommes pour les numéros de ligne suivant, donc mon choix de numéros aléatoires doit respecter ces numéros de ligne : 174, 211, 142, 008, 069, 380, 123, 000, 008, 012, 061, 017, 115, 034, 161, 036, 023, 145, 407, 484, 396

Les âges correspondants : 34, 40, 50, 47, 46, 45, 40, 33, 32, 45, 32, 50, 30, 33, 50, 35, 37, 38, 50, 30, 43, 48, 48, 39, 32, 49, 49, 39, 50, 33

Réalisation de l'estimation ponctuelle : avec la moyenne comme estimateur

On a donc pour moyenne estimée m :

$\bar{X} = 1/50 * \sum xi = 1195/ 30 = 39.83$

Estimation de m par un intervalle de confiance de 1 - alpha

Soit I l'intervalle de confiance pour m au niveau de confiance 1 - α avec $\alpha \in \{10\%, 5\%, 1\%\}$

L'intervalle est sous la forme :

$I = [\bar{X} - k ; \bar{X} + k]$

où k est la marge d'erreur.

- **Détermination de I :**

Nous devons d'abord trouver la valeur de k telle que :

$$P(\bar{X} - k < m < \bar{X} + k) = 1 – \alpha$$

- **Élément de contexte**

X : l'âge des individus

- **Caractérisation de la loi de la variable X et ses paramètres :**

inconnue

- **Tableau :**

	Population : $N = 255$	Échantillon : $n = 30$
Espérance	$E(X) = 39{,}30$	$\bar{X} = 39.83$
Variance		$s2 \approx 37.8$

- **Conditions expérimentales :**

- Population inconnue

- Variance inconnue

- $n > 30$

Construction de I pour un niveau de confiance de 0.05, soit 95 % :

Détermination de k :

$P(\bar{X} - k < m < \bar{X} + k) = 0{,}95$

$\Rightarrow P(-k < m - \bar{X} < k) = 0{,}95$

$\Rightarrow P(-k / (\sigma/\sqrt{n}) < (\bar{X} - m) / (\sigma/\sqrt{n}) < k / (\sigma/\sqrt{n})) = 0{,}95$

Soit $T = (\bar{X} - m) / (\sigma/\sqrt{n}) \sim$ app $N(0,1)$ estimé par $s/\sqrt{n}$ on a $T = (\bar{X} - m) / (s/\sqrt{n})$

$\Rightarrow P(-a < T < a) = 0{,}95$ avec $a = k / (s/\sqrt{n})$

$\Rightarrow F(a) - F(-a) = 0{,}95$

$\Rightarrow 2F(a) = 0{,}95 \Rightarrow F(a) = 0{,}975 \Rightarrow$ car $F(a)=$ **$1 - \alpha/2$**

$a = 1{,}96$ (table de loi normale)

- **Calcul de k :**

$k = 1{,}96 \times (s / \sqrt{n}) = 1{,}96 \times (6.14 / \sqrt{30}) \approx 2.2$

- **Intervalle de confiance à 95 % :**

$I = [39.83 – 2.2; 39.83 + 2.2] = [37.63 ; 42.03]$

Conclusion :

Cela signifie que l'intervalle de confiance à 95% pour la moyenne des âges de la population se situe entre 37.63 et 42.03.

Construction de I pour un niveau de confiance de 0.1, soit 90 % :

$P(\bar{X} - k < m < \bar{X} + k) = 0,9$

$\Rightarrow P(-k < m - \bar{X} < k) = 0,9$

$\Rightarrow P(-k / (\sigma/\sqrt{n}) < (\bar{X} - m) / (\sigma/\sqrt{n}) < k / (\sigma/\sqrt{n})) = 0,9$

Soit $T = (\bar{X} - m) / (\sigma/\sqrt{n}) \sim$ app $N(0,1)$ estimé par $s/\sqrt{n}$ on a $T = (\bar{X} - m) / (s/\sqrt{n})$

$\Rightarrow P(-a < T < a) = 0,9$ avec $a = k / (s/\sqrt{n})$

$\Rightarrow F(a) - F(-a) = 0,9$

$\Rightarrow 2F(a) = 0,9 \Rightarrow F(a) = 0,95 \Rightarrow a = 1,645$ (table de loi normale)

- **Calcul de k :**

$k = 1,645 \times (s / \sqrt{n}) = 1,645 \times (6.14 / \sqrt{30}) \approx 1,796$

- **Intervalle de confiance à 90 % :**

$I = [39.83-1,796, 39.83+1,796]=[38.034, 41,626]$

Conclusion :

Cela signifie que l'intervalle de confiance à 90% pour la moyenne des âges de la population se situe entre 38.034 et 41.626.

Construction de I pour un niveau de confiance de 0.01, soit 99% :

$P(\bar{X} - k < m < \bar{X} + k) = 0,99$

$\Rightarrow P(-k < m - \bar{X} < k) = 0,99$

$\Rightarrow P(-k / (\sigma/\sqrt{n}) < (\bar{X} - m) / (\sigma/\sqrt{n}) < k / (\sigma/\sqrt{n})) = 0,99$

Soit $T = (\bar{X} - m) / (\sigma/\sqrt{n}) \sim$ app $N(0,1)$ estimé par $s/\sqrt{n}$ on a $T = (\bar{X} - m) / (s/\sqrt{n})$

$\Rightarrow P(-a < T < a) = 0,99$ avec $a = k / (s/\sqrt{n})$

$\Rightarrow F(a) - F(-a) = 0,99$

$\Rightarrow 2F(a) = 0,99$

$F(a)= 0,995$

$a=2,576$

- **Calcul de k :**

k = 2,576 × (s / √n) = 2,576 × (6.14/ √30) ≈ 2.89

- **Intervalle de confiance à 99 % :**

I = [39.83−2,89,39.83+2,89]=[36,94,42.72]

Conclusion :

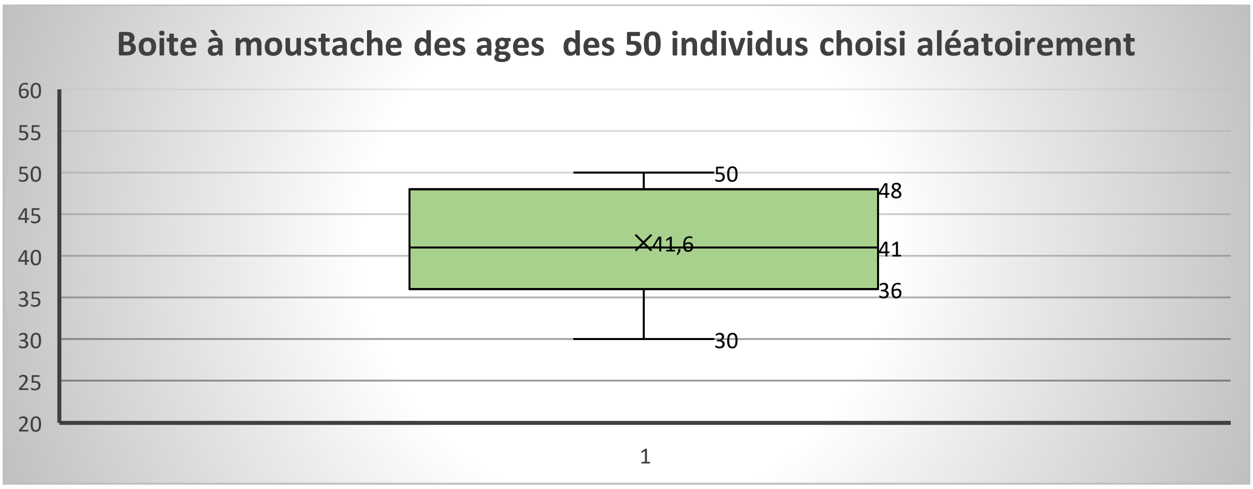
Cela signifie que l'intervalle de confiance à 99% pour la moyenne des âges de la population se situe entre 36.94 et 42.72.

Résumé des intervalles de confiances :

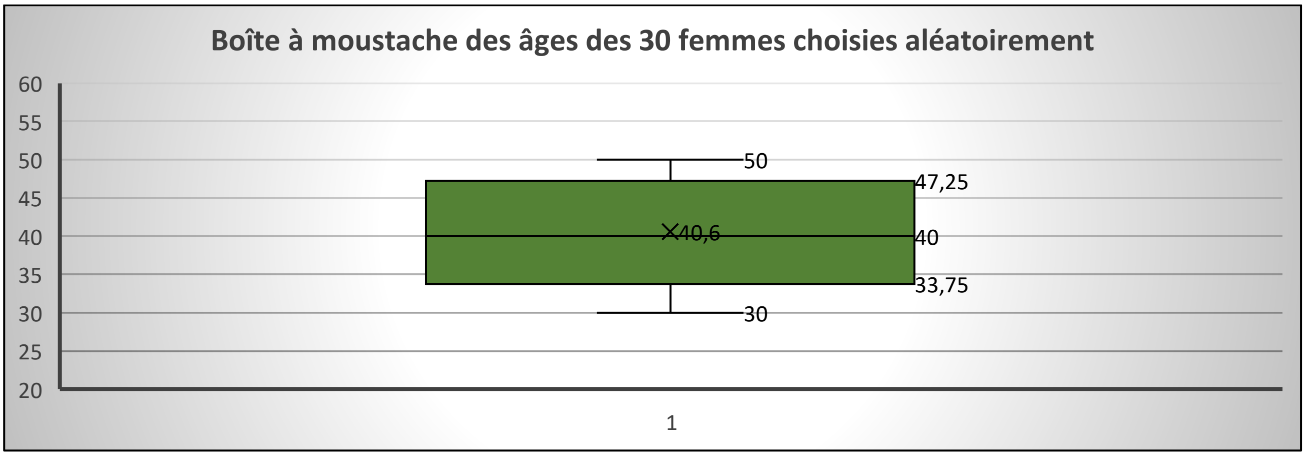
Niveau de confiance	Population totale	Femmes	Hommes
95%	[39,92 ; 43,28]	[40,13 ; 44,41]	[37,63 ; 42,03]
90%	[40,19 ; 43,01]	[40,47 ; 44,07]	[38,03 ; 41,63]
99%	[39,387 ; 43,813]	[39,46 ; 45,08]	[36,94 ; 42,72]

Boites à moustache des échantillons :

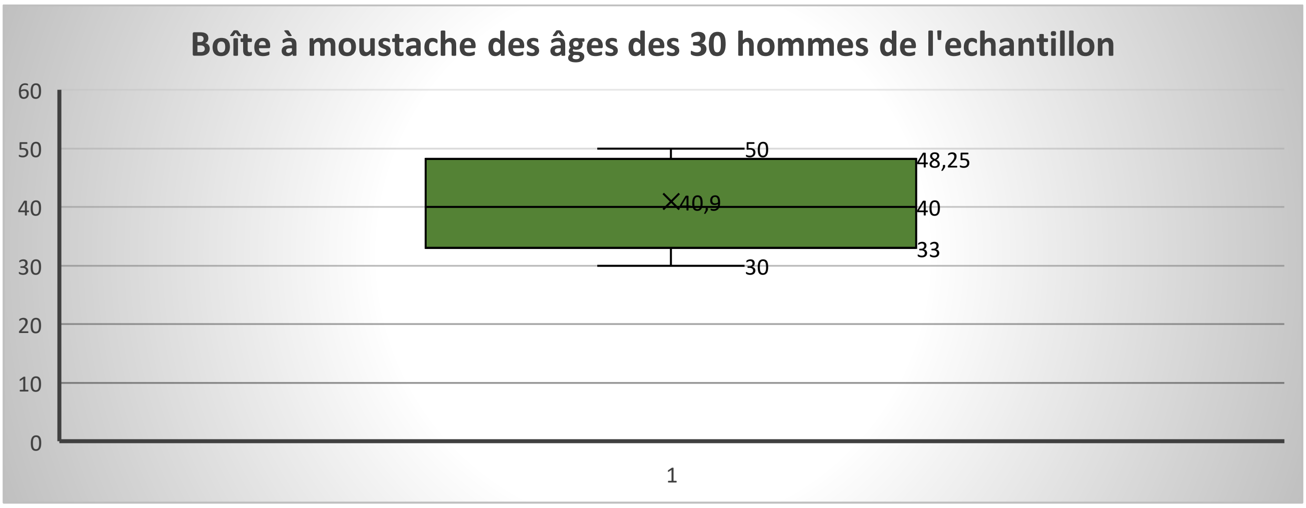
Population totale :



Populations des femmes



Populations des hommes :



Interprétation :

I. Analyse de la moyenne estimée et des intervalles de confiance de l'échantillon de la population totale :

1. Intervalle de confiance à 95 % (niveau de confiance de 0,05)

I=[39,92;43,28]I = [39,92 ; 43,28]I=[39,92;43,28]

Interprétation :
Vous êtes **95 % certain** que la vraie moyenne de la population se situe entre 39,92 et 43,28. La moyenne de la population (39,5) est **en dehors** de cet intervalle, ce qui suggère que l'échantillon n'est pas parfaitement représentatif.

2. Intervalle de confiance à 90 % (niveau de confiance de 0,1)

I=[40,19;43,01]I = [40,19 ; 43,01]I=[40,19;43,01]

Interprétation :
Vous êtes **90 % certain** que la moyenne de la population est entre 40,19 et 43,01. La moyenne de la population (39,5) est encore **en dehors** de cet intervalle, mais l'incertitude est plus faible qu'à 95 %.

3. Intervalle de confiance à 99 % (niveau de confiance de 0,01)

I=[39,387;43,813]I = [39,387 ; 43,813]I=[39,387;43,813]

Interprétation :
Vous êtes **99 % certain** que la vraie moyenne est entre 39,387 et 43,813. La moyenne de la population (39,5) est dans cet intervalle, ce qui suggère que l'échantillon est plus représentatif.

Conclusion :

Les intervalles à 95 % et 90 % ne contiennent pas la moyenne de la population, tandis qu'à 99 %, l'intervalle l'inclut. Cela montre que l'estimation est plus précise à ce niveau de confiance plus élevé.

II. Analyse de la moyenne estimée et des intervalles de confiance de l'échantillon de la population des femmes :

1. Intervalle de confiance à 95 % (niveau de confiance de 0,05)

I=[40,13;44,41]

Interprétation :
On est **95 % certain** que la vraie moyenne d'âge des femmes dans la population se situe entre 40,13 et 44,41.La moyenne réelle (39,7) est **en dehors** de cet intervalle, ce qui suggère que l'échantillon utilisé pourrait ne pas parfaitement représenter l'ensemble des femmes de la population.

2. Intervalle de confiance à 90 % (niveau de confiance de 0,1)

I=[40,474;44,066]

Interprétation :
On est **90 % certain** que la moyenne d'âge des femmes se situe entre 40,47 et 44,07. La moyenne de la population (39,7) est encore **en dehors** de cet intervalle, ce qui suggère que l'échantillon utilisé pourrait ne pas bien représenter l'ensemble des femmes de la population.

3. Intervalle de confiance à 99 % (niveau de confiance de 0,01)

I=[39,458;45,082]

Interprétation :
Vous êtes **99 % certain** que la moyenne d'âge réelle se situe entre 39,46 et 45,08. Cette fois, la moyenne de la population (39,7) est **incluse** dans l'intervalle, ce qui indique que l'estimation est plus fiable et que l'échantillon est surement représentatif.

Conclusion :

Les intervalles à 95 % et 90 % ne contiennent pas la moyenne d'âge réelle des femmes (39,5), mais l'intervalle à 99 % l'inclut, ce qui montre une estimation plus sûre à ce niveau de confiance.

III. Analyse de la moyenne estimée et des intervalles de confiance de l'échantillon de la population des hommes :

1. Intervalle de confiance à 95 % (niveau de confiance de 0,05)

I=[37,63;42,03]

Interprétation :
On est **95 % certain** que la vraie moyenne d'âge des hommes dans la population se situe entre 37,63 et 42,03. La moyenne réelle (39,3) est **dans** cet intervalle, ce qui indique que l'échantillon est représentatif de la population des hommes.

2. Intervalle de confiance à 90 % (niveau de confiance de 0,1)

I=[38,03;41,63]

Interprétation :
On est **90 % certain** que la moyenne d'âge des hommes se situe entre 38,03 et 41,63. La moyenne de la population (39,3) est **incluse** dans cet intervalle, ce qui prouve que l'estimation est plus fiable.

3. Intervalle de confiance à 99 % (niveau de confiance de 0,01)

I=[36,94;42,72]

Interprétation :
Avec un **niveau de confiance de 99 %**, vous êtes presque sûr que la moyenne réelle des âges se situe entre 36,94 et 42,72. La moyenne de 39,3est **bien comprise** dans cet intervalle, ce qui confirme la **très bonne représentativité** de l'échantillon.

Conclusion :

Tous les intervalles (90 %, 95 %, et 99 %) contiennent la moyenne d'âge des hommes dans la population (39,5), ce qui montre que l'échantillon est **représentatif**, peu importe le niveau de confiance choisi.

IV. Analyse globale :

- L'âge **moyen estimé des femmes (42,27 ans)** est **supérieur** à celui des **hommes (39,83 ans)** dans l'échantillonnage. Preuve que les femmes sont plus âgées que les hommes. Ce qui était déjà le cas dans l'analyse de le population total avec 39.7ans pour les femmes et 39.3 pour les hommes.
- Il y'a un léger **chevauchement** entre l'intervalle de la moyenne des âges des hommes et de la moyenne estimé des femmes qui **ne permet pas d'affirmer avec certitude** qu'il y a une **différence** entre l'âge des hommes et des femmes. Par exemple :

  IC à 95 % pour les femmes :
  [40,13;44,41]
  IC à 95 % pour les hommes :
  [37,70;41,96]

  L'**intervalle des femmes commence à 40,13** et celui des hommes **se termine à 41,96**. Il y a donc un **chevauchement** entre **40,13 et 41,96** : les deux intervalles **partagent une plage commune**.

Conclusion générale :

Mon analyse montre que :

- L'âge moyen estimé varie légèrement entre les groupes, avec les femmes présentant une moyenne un peu plus élevée.
- Les intervalles de confiance donnent une estimation **fiable** de l'âge moyen dans chaque population.