ECOLE DES HAUTES ETUDES COMMERCIALES

Module : Statistique de Gestion
Niveau : 1ère année Master (TC)
Groupes : 1, 2, 7 et 8

Année Académique : 2013/2014
Enseignant : KHERRI Abdenacer
Site web : www.sg-ehec.jimdo.com

Support pédagogique de cours N° 02 :

Echantillonnage

Plan du cours :

1. **Introduction** :

L'étude de propriétés caractéristiques d'un ensemble, quand on ne dispose pas encore de données, nécessite d'examiner, d'observer des éléments de cet ensemble. La manière de recueillir ces données fait l'objet d'une théorie mathématique appelée théorie des sondages ou encore théorie de l'échantillonnage (en anglais : *sampling theory*), Cette théorie concerne l'optimisation de la collecte des données selon divers critères et répond à certaines interrogations sur la façon de procéder à cette collecte en rapport avec l'information disponible et l'effort d'échantillonnage consenti.

2. **Terminologie** :

Terme	Définition
Cadre d'échantillonnage	Une liste d'élément à partir desquels l'échantillon est sélectionné.
Distribution d'échantillonnage	Distribution de probabilité composée de toutes les valeurs possibles d'une statistique d'échantillon.
Echantillon	Un sous ensemble de la population considérée, le nombre d'individus dans l'échantillon est la taille de l'échantillon.
Echantillonnage	La sélection d'une partie dans un tout (la sélection d'une partie dans la population), l'échantillon sélectionné doit être représentatif de la population.
Méthodes d'échantillonnage	Ensemble des méthodes permettant de réaliser un sondage (de prélever un échantillon de données) au sein d'une population, de manière à reproduire un échantillon aussi représentatif que possible de cette population.
Paramètre	Caractéristique numérique d'une population telle que la moyenne de la population "μ", l'écart type de la population "σ" et la proportion de la population "p".
Population	Ensemble que l'on observe et qui sera soumis à une analyse statistique, chaque élément de cet ensemble est un individu ou unité statistique.
Population finie	Une population qui consiste en un nombre fini d'éléments.
Population infinie	Une population est infinie s'il n'y a pas de limite au nombre d'éléments qu'il contient.
Population homogène	Une population avec des éléments qui possèdent les mêmes caractéristiques.
Population non homogène	Une population avec des éléments qui ne possèdent pas les mêmes caractéristiques.
Tirage exhaustif	Tirage sans remise.
Tirage non exhaustif	Tirage avec remise.
Population échantillonnée	La population à partir de laquelle l'échantillon est constitué.
Théorie d'échantillonnage	Etude des liaisons existant entre une population et les échantillons de cette population.

3. **Echantillonnage** :

Définition (1) : L'échantillonnage est le procédé utilisé pour choisir un échantillon et qui est à la base de l'enquête par sondage[1].

Définition (2) : l'échantillonnage est la phase qui consiste à sélectionner les individus que l'on souhaite interroger au sein de la population de base.[2]

Prenons tous les échantillons possibles de taille **n** tirés d'une population donnée. Pour chaque échantillon, on peut calculer une statistique (moyenne, écart-type, variance, etc...) qui variera avec l'échantillon. Pour tous les échantillons, on obtient alors une distribution de la statistique que l'on nomme *la distribution d'échantillonnage*. Pour la validité des résultats, il est important que les échantillons soient représentatifs de la population concernée.

Combien d'échantillons de **n** éléments peuvent être isolés d'une population de **N** éléments ?

On distingue entre deux cas de tirage :

- Tirage exhaustif (sans remise) : nombre d'échantillons est C_N^n
- Tirage non exhaustif (avec remise) : nombre d'échantillons est N^n

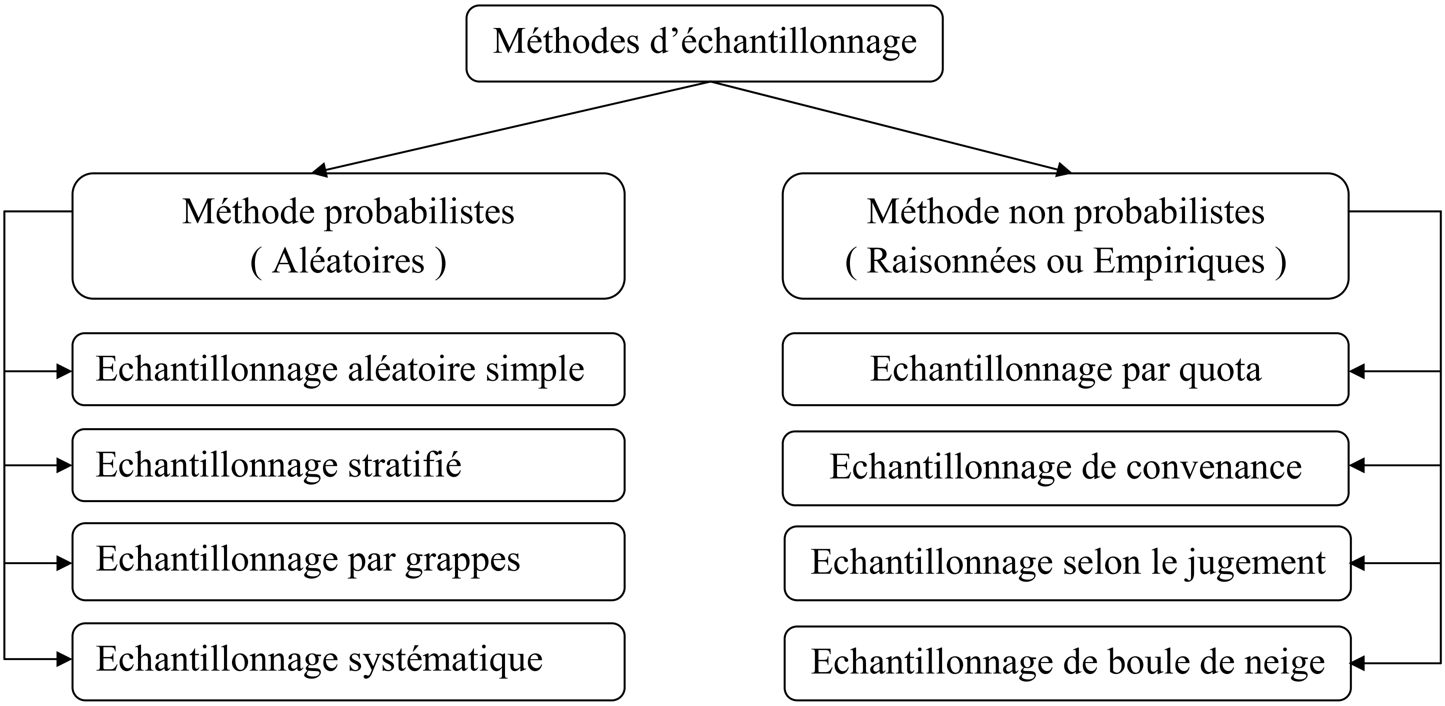
3.1. **Méthodes d'échantillonnage** :

L'échantillonnage peut se faire avec ou sans remise et une population peut être considérée comme finie ou infinie. Une population finie dans laquelle on procède à un échantillonnage avec remise peut être théoriquement considérée comme infinie.

Dans la pratique, il en va de même pour des populations finies mais de grandes tailles. Pour chaque distribution d'échantillonnage, on peut calculer une moyenne, un écart type, une variance…etc.

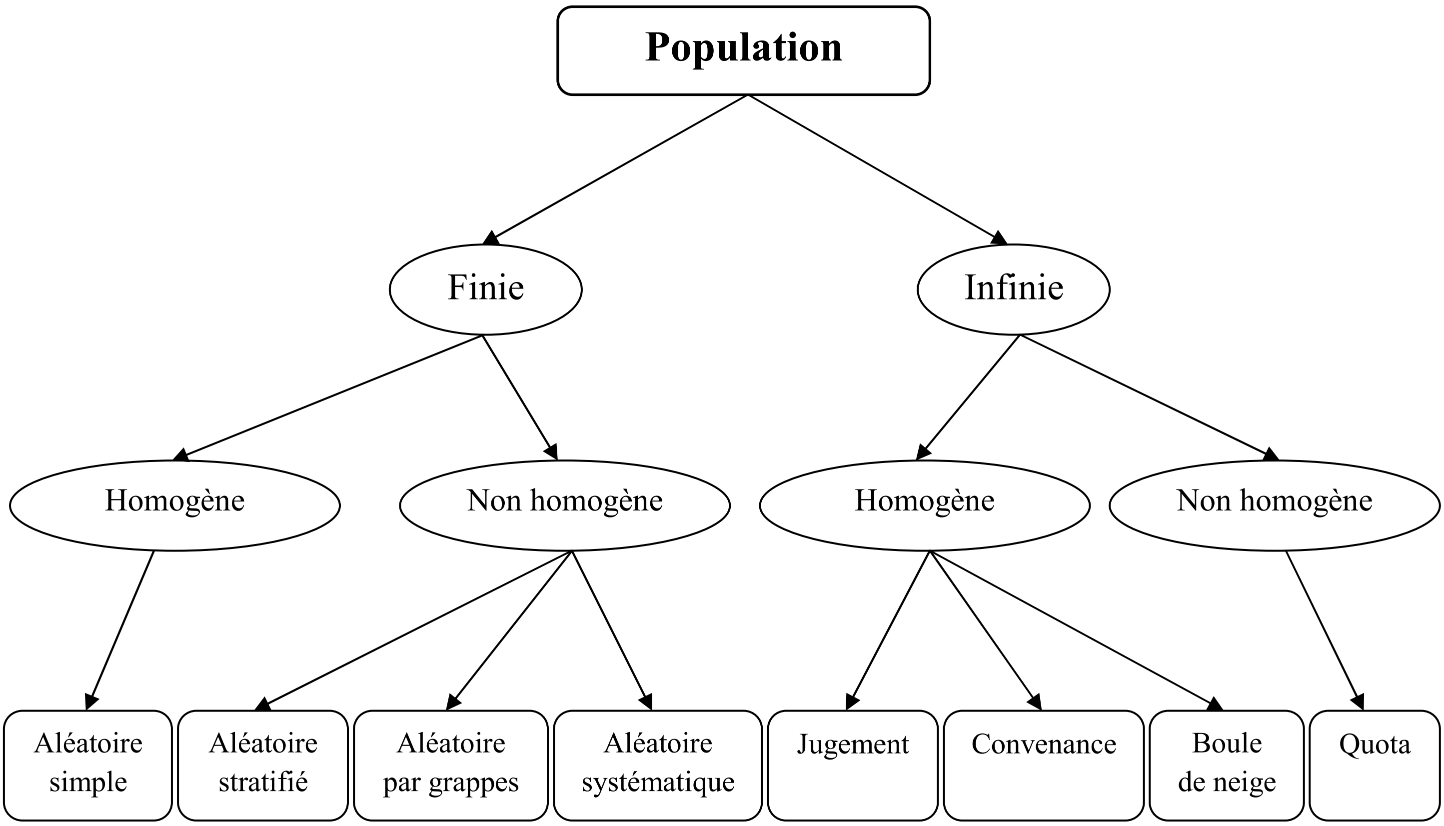
1. http://www.larousse.fr/dictionnaires/francais/echantillonnage
2. http://www.definitions-marketing.com/Definition-Echantillonnage-etude

Schéma N° 01 : Les méthodes d'échantillonnage



Source : élaboré par l'enseignant

Schéma N° 02 : Les méthodes d'échantillonnage



<u>**المصدر**</u> : محمد حسين محمد رشيد (القادري) و منى عطا الله الشويلات، <u>**مبادئ الإحصاء و الاحتمالات و معالجتها باستخدام برنامج (SPSS)**</u>، الطبعة الأولى، دار صفاء للنشر و التوزيع، عمان، الأردن، 2012، الصفحة 28.

3.1.1. <u>Méthodes probabilistes (Aléatoires)</u> :

L'échantillonnage probabiliste repose sur un choix d'unités dans la population fait au hasard, ce n'est pas l'enquêteur qui choisit les unités, c'est la méthode utilisée pour la sélection qui le fait. Une des caractéristiques de cette méthode est que chaque unité de la population a une probabilité mesurable d'être choisie.

L'avantage de la méthode d'échantillonnage probabiliste est qu'elle permet de généraliser les résultats de l'échantillon à l'ensemble de la population en s'appuyant sur une théorie statistique reconnue.

Son seul inconvénient est qu'il faut posséder une liste de toutes les unités formant la population avant de procéder à la sélection de l'échantillon.

Voici les quatre types d'échantillonnage probabiliste que l'on peut effectuer :

3.1.1.1. <u>Echantillonnage aléatoire simple</u> :

Un échantillon aléatoire simple est un échantillon sélectionnée de manière à ce que chaque échantillon possible de taille "**n**" ait la même probabilité d'être sélectionné, On prélève dans la population des individus au hasard, tous les individus ont la même probabilité d'être prélevés, et ils le sont indépendamment les uns des autres.

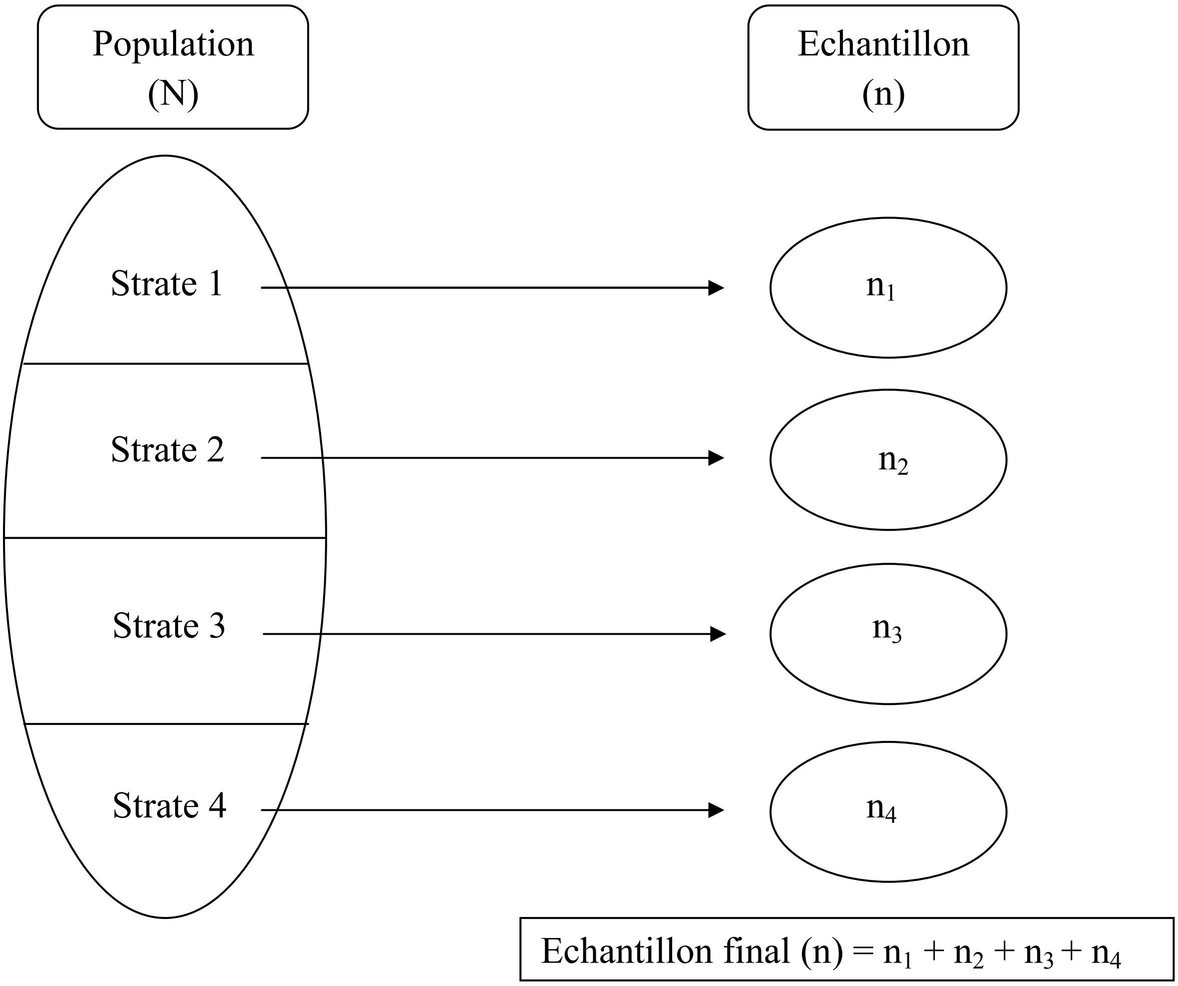
3.1.1.2. <u>Echantillonnage aléatoire stratifié</u> :

On suppose que la population soit stratifiée, constituée de sous-populations homogènes, les strates. (ex : stratification par tranche d'age). Dans chaque strate, on fait un échantillonnage aléatoire simple, de taille proportionnelle à la taille de strate dans la population (échantillon représentatif). Les individus de la population n'ont pas tous la même probabilité d'être tirés. Nécessite une homogénéité des strates.

Le chercheur divise la population en sous-groupes distincts et homogènes (strates) à partir desquels il sélectionnera un échantillon aléatoire simple.

<u>Étapes :</u>

1. choisir une variable de stratification (ex : tranche d'age).
2. Sélectionner un échantillon aléatoire dans chaque strate.



<u>Exemple</u> :

Supposons que 60% des étudiants de l'école HEC sont des filles et 40% des garçons, pour former un échantillon de 120 étudiants en respectant ces strates, on devrait choisir au hasard 60% x 120 = 72 filles et 40% x 120 = 48 garçons.

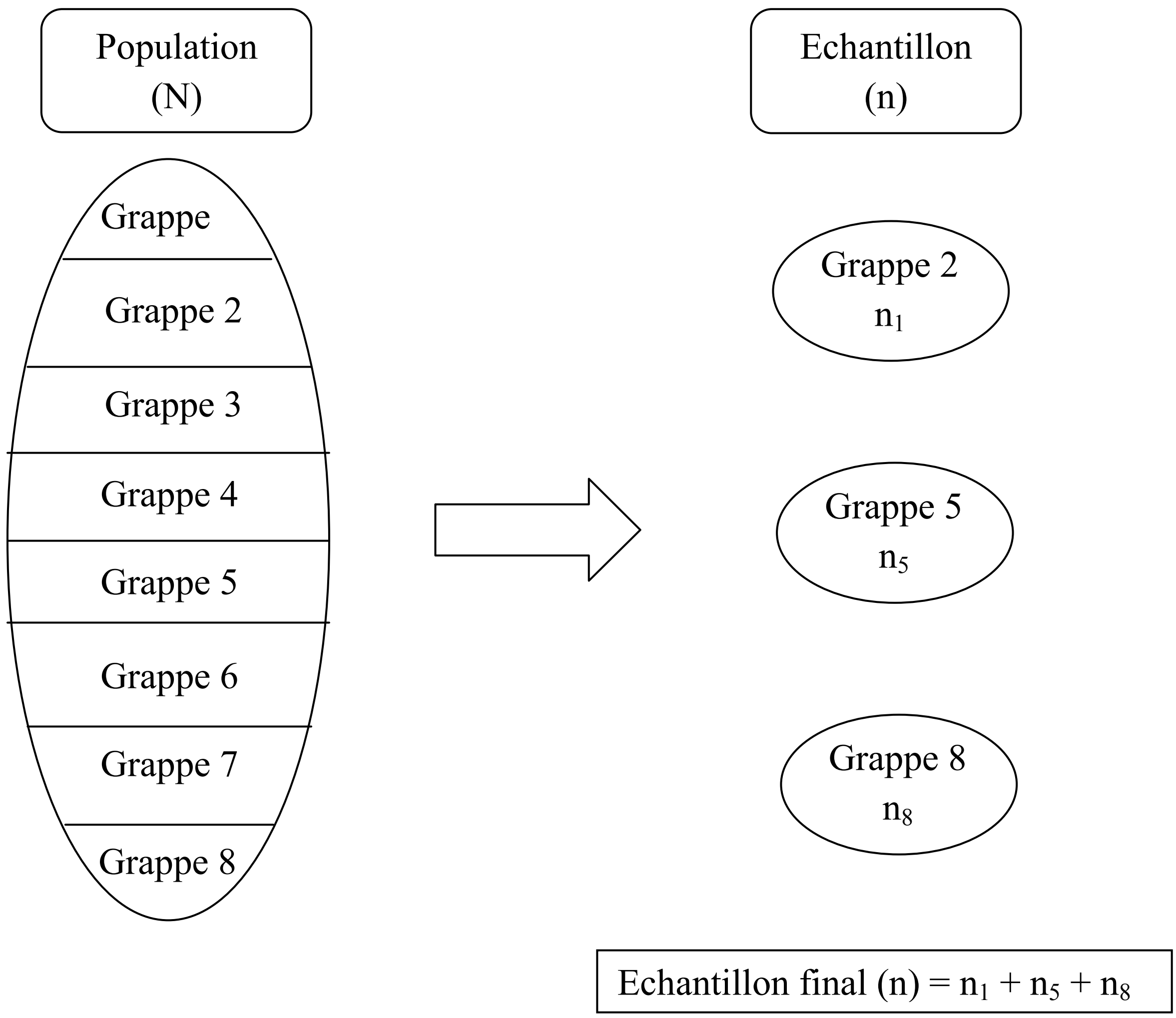
<u>Avantages et désavantages de la méthode</u> :

L'échantillonnage stratifié a l'avantage d'assurer une bonne représentation des différentes strates de la population dans l'échantillon. Il permet aussi d'obtenir des estimations pour chacune des strates de la population. Toutefois, pour utiliser cette méthode il faut avoir des renseignements sur la répartition des strates dans la population.

3.1.1.3. **Echantillonnage aléatoire par grappe :**

On tire au hasard des grappes ou familles d'individus, et on examine tous les individus de la grappe (ex: on tire des immeubles puis on interroge tous les habitants). La méthode est d'autant meilleure que les grappes se ressemblent et que les individus d'une même grappe sont différents, contrairement aux strates.

Le chercheur divise la population en sous-groupes appelés « grappes ». Les grappes ont le même profil, la variance d'une grappe à l'autre étant faible. Il sélectionne par la suite un échantillon aléatoire de grappes et non pas un échantillon aléatoire à l'intérieur de chaque grappe.



Exemple :

Les étudiants de première année Master à HEC sont répartis en 11 groupes, les groupes sont numérotés de 1 à 11. Supposons que l'on obtienne les nombres 2, 5, 7 et 10, tous les étudiants de ces 4 groupes feront partie de l'échantillon.

Avantages et désavantages de la méthode :

L'avantage de cette méthode par rapport aux précédentes est qu'elle ne requiert pas au préalable la liste de la population, seule la liste des unités pour les grappes pigées est nécessaire. Un désavantage de ce type d'échantillonnage est qu'il produit des estimations habituellement moins précises que l'échantillonnage aléatoire simple parce que des unités appartenant à une même grappe ont tendance à présenter des caractéristiques semblables. Cette perte de précision peut être compensée par une augmentation de la taille de l'échantillon.

3.1.1.4. Echantillonnage aléatoire systématique :

Dans certaines situations, spécialement lorsque les populations sont importantes, il est coûteux (en temps) de sélectionner un échantillon aléatoire simple en trouvant tout d'abord un nombre aléatoire et ensuite en cherchant dans la liste de la population l'élément correspondant. Une alternative de l'échantillonnage aléatoire simple est ***l'échantillonnage systématique***. Par exemple, si l'on souhaite sélectionner un échantillon de taille **50** parmi une population contenant **5000** éléments, cela revient à sélectionner un élément tous les **(5000/50) = 100** éléments de la population. Constituer un échantillon systématique dans ce cas consiste à sélectionner aléatoirement un élément parmi les 100 premiers de la liste de la population. Les autres éléments de l'échantillon sont identifiés de la façon suivante : le second élément sélectionné correspond au 100^{e} élément qui suit le premier élément sélectionné dans la liste de la population, le troisième élément sélectionné correspond au 100^{e} élément qui suit le deuxième élément sélectionné dans la liste de la population, et ainsi de suite. En fait, l'échantillon de taille 50 est identifié en se déplaçant systématiquement dans la population et en identifiant les 100^{e}, 200^{e}, 300^{e} …etc. éléments qui suivent le premier élément choisi aléatoirement. L'échantillon de taille 50 est généralement plus facile à identifier de cette manière qu'en utilisant l'échantillonnage aléatoire simple. Puisque le premier élément sélectionné l'est aléatoirement, un échantillon systématique est généralement supposé avoir les propriétés d'un échantillon aléatoire simple, cette hypothèse est particulièrement appropriée lorsque la liste de la population est une énumération aléatoire des éléments de la population[1].

1. Anderson, Sweeney et Williams, ***Statistique pour l'économie et la gestion***, édition De Boeck, Bruxelles, Belgique, 2010, P364.

3.1.2. **Méthodes non probabilistes (Raisonnées ou empirique) :**

L'échantillonnage non probabiliste repose sur un choix arbitraire des unités, c'est l'enquêteur qui choisit les unités et non le hasard. En ce sens, il serait donc aventureux de généraliser les résultats obtenus pour l'échantillon à toute la population. Malgré cela, ces méthodes sont souvent utilisées dans certaines disciplines. En voici quelques-unes :

3.1.2.1. **Echantillonnage par quota :**

Lorsque le chercheur veut reproduire les caractéristiques d'une population (ex. âge, sexe, revenus, etc.) dans son échantillon.

3.1.2.2. **Echantillonnage de convenance (de commodité) :**

Cas où les unités d'échantillonnage sont faciles à rejoindre, disponibles et généralement facile à convaincre.

3.1.2.3. **Echantillonnage selon le jugement :**

Le chercheur juge que l'échantillon va lui permettre d'atteindre les objectifs de la recherche.

3.1.2.4. **Echantillonnage boule de neige :**

Utile dans le cas de la rareté des unités d'échantillonnage ou de l'absence d'un cadre d'échantillonnage valide. On demande à un répondant de nous référer à un autre qui présente les mêmes caractéristiques que les siennes, et ainsi de suite…

4. Distribution d'échantillonnage :

La distribution d'échantillonnage est l'étude de la de probabilité de l'échantillon en fonction de la distribution de la variable parente lorsque la taille de l'échantillon augmente.

Pour résoudre les problèmes d'estimation de paramètres inconnus, il faut tout d'abord étudier les ***distributions d'échantillonnage***, c'est à dire la loi de probabilité suivie par l'estimateur.

Considérons tous les échantillons possibles de taille n extraits d'une population de taille N, de moyenne μ, de variance σ^2, …etc. Pour chaque échantillon, il est possible de calculer les paramètres statistiques $\bar{x}$, s, s^2…etc qui varient d'un échantillon à l'autre. Chaque paramètre possédera ainsi une distribution d'échantillonnage au même titre que la variable aléatoire **X**.

On utilise souvent :

	Population	Echantillon	Distribution d'échantillonnage des moyennes	Distribution d'échantillonnage des variances
Taille	N	n	/	/
Moyenne	μ	$\bar{x}$	$\mu_{\bar{X}}$	μ_{s^2}
Ecart-type	σ	s	$\sigma_{\bar{X}}$	σ_{s^2}
Variance	σ^2	s^2	$\sigma^2{}_{\bar{X}}$	$\sigma^2{}_{s^2}$
Proportion	π	p	/	/

4.1. Distribution d'échantillonnage des moyennes :

Soit une population de taille **N**, on désigne par μ et σ la moyenne et l'écart-type de cette population respectivement. On extrait de la population une série d'échantillons de taille **n**, chacun de ces échantillons a une moyenne $\bar{x}$, les différentes moyennes obtenues constituent une distribution d'échantillonnage de moyenne $\bar{X}$, on désigne par $\mu_{\bar{X}}$ et $\sigma_{\bar{X}}$ la moyenne et l'écart-type de la distribution d'échantillonnage de la moyenne.

On a : $\mu_{\overline{X}} = \mu$

Si le tirage est exhaustif (sans remise) : $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$

Dans le cas où la population est infinie ou le tirage est non exhaustif (avec remise) : $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$

Remarques[1] :

- Si **n** est petit devant **N**, la distinction entre exhaustivité et non exhaustivité est sans objet car $\frac{N-n}{N-1} \approx 1$
- Si la taille des échantillons est assez grande (en pratique $n \geq 30$), la distribution d'échantillonnage de la moyenne approche la distribution normale quelle que soit la distribution de la population.
- Si la population est normalement distribuée, la distribution d'échantillonnage de la moyenne est une loi normale quelle que soit la valeur **n** de la taille des échantillons.

Exemple corrigé : [exemple pour but de confirmer les formules précédentes]

On a une population finie composée de 3 éléments,

$P = \{1, 2, 3\}$

- Calculer la moyenne et l'écart-type de cette population.

Réponse :

On sait que $\mu = \frac{\sum_{i=1}^{n} x_i}{N} = \frac{1+2+3}{3} = \frac{6}{3} = 2$

Et on sait aussi que $\sigma = \sqrt{\sigma^2}$ *et que* $\sigma^2 = \frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{n} = \frac{\sum_{i=1}^{n} {x_i}^2}{n} - \overline{x}^2$

Donc : $\sigma^2 = \frac{\sum_{i=1}^{n} {x_i}^2}{n} - \overline{x}^2 = \frac{1+4+9}{3} - 4 = \frac{14}{3} - 4 = \frac{14-12}{3} = \frac{2}{3}$

$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{2}{3}}$

Résultats finaux : $\mu = 2$ *et* $\sigma = \sqrt{\frac{2}{3}}$

1. KHALDI Khaled, ***Méthodes statistiques (rappel de cours et exercices corrigés)***, 6ème édition, OPU, Alger, Algérie, 2010, P76.

On va effectuer des prélèvements des échantillons de taille ($\boldsymbol{n = 2}$), on fait le prélèvement dans les deux cas (tirage exhaustif et tirage non exhaustif).

Cas 1 (Tirage non exhaustif / tirage avec remise) :

- Quelle est le nombre des échantillons qui peuvent être prélevés à partir de cette population ?
- Effectuez le prélèvement de ces échantillons.
- Etablir une distribution d'échantillonnage des moyennes.
- Calculez la moyenne de la distribution d'échantillonnage des moyenne $\boldsymbol{\mu_{\overline{X}}}$.
- Calculez l'écart-type de la distribution d'échantillonnage des moyenne $\boldsymbol{\sigma_{\overline{X}}}$.
- Peut-on nous confirmer les deux formules suivantes $\boldsymbol{\mu_{\overline{X}} = \mu}$ et $\boldsymbol{\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}}$?

Réponse :

- *Parce que le tirage est avec remise donc le nombre des échantillons possible à être prélevés est une liste n éléments pris parmi N éléments c-à-d* $\boldsymbol{N^n = 3^2 = 9}$
- *Les échantillons sont les suivants :*

$$E = \begin{Bmatrix} (1,1)(1,2)(1,3) \\ (2,1)(2,2)(2,3) \\ (3,1)(3,2)(3,3) \end{Bmatrix}$$

- *La distribution d'échantillonnage des moyennes est la suivante :*

$$\textit{La distribution d'échantillonnage des moyennes} = \begin{Bmatrix} (1)(1,5)(2) \\ (1,5)(2)(2,5) \\ (2)(2,5)(3) \end{Bmatrix}$$

- *La moyenne de la distribution d'échantillonnage des moyennes :*

$$\mu_{\overline{X}} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{18}{9} = 2 = \mu$$

- *L'écart-type de la distribution d'échantillonnage des moyennes :*

$$\sigma_{\overline{X}}{}^2 = \frac{\sum_{i=1}^{n} x_i{}^2}{n} - \overline{x}^2 = \frac{39}{9} - 4 = \frac{39-36}{9} = \frac{3}{9} = \frac{1}{3}$$

$$\sigma_{\overline{X}} = \sqrt{\sigma_{\overline{X}}{}^2} = \sqrt{\frac{1}{3}}$$

- *Confirmation :*

$\boldsymbol{\mu_{\overline{X}} = 2 = \mu}$ *formule confirmée*

$\boldsymbol{\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{\sqrt{\frac{2}{3}}}{\sqrt{2}} = \sqrt{\frac{1}{3}}}$ *formule confirmée*

Cas 2 (Tirage exhaustif / tirage sans remise) :

- Quelle est le nombre des échantillons qui peuvent être prélevés à partir de cette population ?
- Effectuez le prélèvement de ces échantillons.
- Etablir une distribution d'échantillonnage des moyennes.
- Calculez la moyenne de la distribution d'échantillonnage des moyenne $\boldsymbol{\mu_{\overline{X}}}$.
- Calculez l'écart-type de la distribution d'échantillonnage des moyenne $\boldsymbol{\sigma_{\overline{X}}}$.
- Peut-on nous confirmer les deux formules suivantes $\boldsymbol{\mu_{\overline{X}} = \mu}$ et $\boldsymbol{\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}}$?

Réponse :

- *Parce que le tirage est sans remise donc le nombre des échantillons possible à être prélevés est une combinaison de n éléments pris parmi N éléments c-à-d* $C_3^2 = \frac{3!}{2!(3-2)!} = \frac{6}{2} = 3$
- *Les échantillons sont les suivants :*

$$\boldsymbol{E = \{(1,2)(1,3)(2,3)\}}$$

- *La distribution d'échantillonnage des moyennes est la suivante :*

  $\boldsymbol{La\ distribution\ d'échantillonnage\ des\ moyennes = \{(1,5)(2)(2,5)\}}$
- *La moyenne de la distribution d'échantillonnage des moyennes :*

$$\boldsymbol{\mu_{\overline{X}} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{6}{3} = 2 = \mu}$$

- *L'écart-type de la distribution d'échantillonnage des moyennes :*

$$\boldsymbol{\sigma_{\overline{X}}{}^2 = \frac{\sum_{i=1}^{n} x_i{}^2}{n} - \overline{x}^2 = \frac{12,5}{3} - 4 = \frac{12,5 - 12}{3} = \frac{0,5}{3} = \frac{1}{6}}$$

$$\boldsymbol{\sigma_{\overline{X}}} = \sqrt{\sigma_{\overline{X}}{}^2} = \sqrt{\frac{1}{6}}$$

- *Confirmation :*

  $\boldsymbol{\mu_{\overline{X}} = 2 = \mu}$ *formule confirmée*

  $\boldsymbol{\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}} = \frac{\sqrt{\frac{2}{3}}}{\sqrt{2}} \boldsymbol{\sqrt{\frac{1}{2}}} = \sqrt{\frac{1}{3}} \boldsymbol{\sqrt{\frac{1}{2}}} = \sqrt{\frac{1}{6}}$ *formule confirmé*

Exemples à corriger :

Exemple (01) : [Distribution d'échantillonnage de la moyenne]

On choisit au hasard sans remise six nombres parmi les nombres entiers de **1** à **9**, chacun de ces nombres a la même probabilité d'être choisi[1].

- Quel est le type d'échantillonnage dans ce cas ?
- Quel est le type de tirage dans ce cas ?
- Quel est le nombre des échantillons possible d'être prélever de la population mère ?
- Calculer la moyenne et l'écart-type de la distribution d'échantillonnage des moyennes.

Réponse :

……

……

……

……

……

……

……

……

……

……

……

……

……

……

……

……

……

……

……

1. KHALDI Khaled, ***Méthodes statistiques (rappel de cours et exercices corrigés)***, op.cit, P84.

Exemple (02) : [Distribution d'échantillonnage de la moyenne]

On choisit au hasard avec remise six nombres parmi les nombres entiers de **1** à **9**, chacun de ces nombres a la même probabilité d'être choisi[1].

- Quel est le type d'échantillonnage dans ce cas ?
- Quel est le type de tirage dans ce cas ?
- Quel est le nombre des échantillons possible d'être prélever de la population mère ?
- Calculer la moyenne et l'écart-type de la distribution d'échantillonnage des moyennes.

✍ **Réponse :**

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

1. KHALDI Khaled, ***Méthodes statistiques (rappel de cours et exercices corrigés)***, op.cit, P76.

<u>Exemple (03)</u> : [Distribution d'échantillonnage de la moyenne]

La moyenne des notes d'une épreuve de statistique de **300** étudiants est égale à **9,8** et l'écart-type est de **3,68**. Trouver la probabilité qu'un échantillon aléatoire de notes de **40** étudiants extrait de l'ensemble ait une moyenne :

1. Comprise entre **10** et **13**.
2. Inférieure à **10**.

NB : Trouver la probabilité dans les deux cas (tirage exhaustif et non exhaustif)

<u>Réponse</u> :

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

4.2. <u>Distribution d'échantillonnage des variances</u> :

Chaque échantillon de taille **n** de la population à une variance $S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$, ces variances sont des valeurs observées d'une même variable aléatoire.

$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ On a :

$E(S^2) = \frac{n-1}{n}\sigma^2$

$$V(S^2) = \frac{n-1}{n^3}[(n-1)\mu^4 - (n-3)\sigma^4]$$

4.3. <u>Distribution d'échantillonnage des fréquences</u> :

La probabilité de la réalisation d'un évènement est supposée être égale à p. on considère les échantillons de taille n extraits, avec remise, d'une population de taille N. a chaque échantillon extrait correspond une fréquence f_n de réalisation de l'évènement considéré.

On a :

Si le tirage est avec remise :

$E(f_n) = p$

$$V(f_n) = \frac{p(1-p)}{n}$$

Si le tirage est sans remise :

$E(f_n) = p$

$$V(f_n) = \frac{p(1-p)}{n} . \frac{N-n}{N-1}$$

<u>Exemples</u> :

<u>Exemple (01)</u> : [Distribution d'échantillonnage de la fréquence]

Un fabricant de clous a déterminé que **3%** des clous produits sont défectueux. On étudie un échantillon aléatoire de **300** clous. Quelle est la probabilité que la proportion de clous défectueux dans l'échantillon soit comprise entre **2%** et **3,5**% ?

✍ **Réponse :**

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

4.4. Distribution d’échantillonnage des différences des moyennes :

On considère 2 populations P_1 et P_2 de moyennes μ_1 et μ_2 et de variances σ_1^2 et σ_2^2. On s’intéresse à la différence μ_1 - μ_2.

on a :

$$\mu_{\bar{x}_1-\bar{x}_2} = \mu_1 - \mu_2$$

$$\sigma^2_{\bar{x}_1-\bar{x}_2} = \sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2}$$

$$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2}}$$

<u>Exemple (01)</u> : [Distribution d’échantillonnage des différences des moyennes]

La résistance à la rupture du hêtre et du bouleau sont respectivement de **4500 kg** et de **4000 kg** avec des écarts-type respectifs de **200 kg** et **300 kg**. Si l’on teste des échantillons de **100** bouleaux et **50** hêtres :

- Calculez $\mu_{\bar{x}_1-\bar{x}_2}$ et $\sigma_{\bar{x}_1-\bar{x}_2}$

<u>Réponse</u> :

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

5. **Synthèse** :

	Echantillonnage aléatoire simple	
	Avec remise	Sans remise
Nombre d'échantillons	N^n	C_N^n
Distribution d'échantillonnage de la moyenne	$\mu_{\bar{X}} = \mu$ $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$	$\mu_{\bar{X}} = \mu$ $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$
Distribution d'échantillonnage de la variance	σ^2	$\frac{n}{n-1}S^2$
Distribution d'échantillonnage de l'écart-type	σ	$\sqrt{\frac{n-1}{n}}\sigma$
Distribution d'échantillonnage de la fréquence	$E(f_n) = p$ $V(f_n) = \frac{p(1-p)}{n}$	$E(f_n) = p$ $V(f_n) = \frac{p(1-p)}{n}$

Le tableau ci-dessous envisage tous les cas auxquels vous pouvez être confrontés[1] :

Paramètre d'échantillon à contrôler	Loi de la population	Statistique	Loi
Moyenne $\bar{X}$	Normale ou quelconque avec n > 30	$\sqrt{n}\left(\frac{\bar{X}-\mu}{\sigma}\right)$	N(0,1) ou ~ N(0,1)
Variance S^2	Normale	$\frac{(n-1)S^2}{\sigma^2}$	χ^2 à n-1 d.d.l.
Proportion f	n > 50	$\sqrt{n}\frac{F-p}{\sqrt{p(1-p)}}$	~ N(0,1)

[1] . http://www.iutbayonne.univ-pau.fr/~grau/2A/stat/cadre2.html