

TIRAGE 2.1

Un logiciel destiné à la réalisation de tirages aléatoires

Manuel de l'utilisateur

par Julien Amegandjin

Octobre 2009

Sommaire

Avant-propos.....	2
<u>Chapitre 1 : Présentation générale du logiciel</u>	
1. Introduction	3
2. A propos de l'installation du logiciel	4
3. Principe de la génération des nombres aléatoires et des tirages aléatoires.....	4
4. Les différents modes de tirage d'échantillons	4
5. Les différents groupes de modules du logiciel.....	5
6. La barre des menus du logiciel.....	6
<u>Chapitre 2 : Comment s'exécutent les modules du logiciel ?</u>	
1. Exemple 1 : Cumuler les tailles des unités d'une population.....	11
2. Exemple 2 : Génération selon le mode PEAR d'un échantillon de nombres entiers compris entre M et Z inclus.....	13
3. Exemple 3 : Tirage selon le mode PIAR d'un échantillon d'unités dans une population	14
4. Exemple 4 : Tirages dans un sondage à deux degrés	18
5. Exemple 5 : Génération PEAR de nombres de K chiffres de type texte	26
6. Exemple 6 : Génération PESR d'un échantillon de nombres entiers de K chiffres	28
7. Exemple 7 : Attribuer un numéro d'ordre aux unités d'une population.....	30
8. Exemple 8 : Réordonner au hasard les unités d'une population.....	32
<u>Chapitre 3 : Fondements des méthodes de tirage utilisées</u>	
1. Principe de la génération des nombres aléatoires.....	34
2. Génération PEAR d'un échantillon de nombres de K chiffres de type texte	36
3. Génération PEAR d'un échantillon de nombres entiers compris entre M et Z inclus.....	36
4. Principe des tirages aléatoires	36
5. Tirage PIAR d'un échantillon d'unités statistiques au sein d'une population.....	36
6. Tirage SPE (tirage systématique avec probabilités égales) d'un échantillon d'unités	38
7. Tirage systématique avec probabilités proportionnelles à la taille des unités : une version simplifiée de la procédure de Hartley et Rao	40
8. Tirage systématique avec probabilités proportionnelles à la taille des unités : la version originale de la procédure de Hartley et Rao	43
9. Procédure de tirage SPPT basée sur les probabilités d'inclusion d'ordre 1	44

Avant-propos

Le logiciel TIRAGE 2.1 succède au logiciel TIRAGE 1.0 paru le 23 octobre 2005. Bien qu'il y ait eu une version préliminaire dénommée TIRAGE 2.0 qui a servi de test de mai à octobre 2009, la véritable deuxième version du logiciel TIRAGE est cette version dénommée TIRAGE 2.1 objet du présent manuel de l'utilisateur.

Les principaux réaménagements contenus dans la version TIRAGE 2.1 sont dictés par un besoin de simplification ressenti pour certains modules du logiciel. Sont notamment concernés, les modules relatifs aux tirages PEAR et aux tirages PESR d'un échantillon d'unités au sein d'une population. En outre, deux nouveaux modules ont été ajoutés aux utilitaires du logiciel.

Malgré les réaménagements apportés, le logiciel TIRAGE 2.1 possède le même objet que son prédécesseur. C'est un logiciel destiné à la réalisation de tirages aléatoires. Il permet notamment, la génération de nombres aléatoires, le tirage d'échantillons aléatoires au sein d'une population finie selon différents modes de sélection ou la simulation de processus aléatoires ou phénomènes faisant appel au hasard. Il se présente comme un outil commode d'utilisation pour suppléer l'emploi des tables de nombres au hasard peu pratiques pour des tirages importants.

Le logiciel TIRAGE 2.1 s'adresse à tous ceux qui sont amenés à se servir des tables de nombres au hasard à défaut de disposer d'applications pour résoudre les problèmes de substitution au hasard. Il s'adresse notamment à des statisticiens pour le tirage des échantillons dans le cadre d'enquêtes par sondage et à d'autres utilisateurs de générations de nombres aléatoires aux fins d'exercices de simulation.

Bien qu'ayant fait l'objet d'un dépôt légal et bénéficiant d'une protection, TIRAGE 2.1 est distribué gratuitement. Il ne peut donc faire l'objet d'aucune transaction commerciale, qu'il s'agisse d'une vente ou de toute autre opération lucrative. Je l'offre aux générations des statisticiens plus jeunes en reconnaissance de ce que la vie m'a apporté dans le domaine du savoir. C'est le prolongement naturel de mes actions antérieures de partage du savoir accomplies à travers des cours photocopiés produits et offerts à des étudiants de 1968 à 1985, ou à travers des documents techniques produits en faveur de la formation dans des projets que j'ai dirigés en Afrique subsaharienne de 1987 à 1999 pour le compte de la FAO ou dans le cadre d'ateliers de formation sur les sondage que j'anime depuis la parution de la première version de ce logiciel.

Des utilisateurs qui auraient des observations ou suggestions à formuler sur TIRAGE 2.1 sont aimablement invités à me les faire parvenir. Qu'ils en soient remerciés d'avance.

Fait à Lomé, le 31 Octobre 2009

J. Amegandjin

Chapitre 1

Présentation générale du logiciel

1. Introduction

Le logiciel TIRAGE 2.1 est l'œuvre de Julien Amegandjin¹, un spécialiste des sondages et un habitué des tirages aléatoires. Il est développé avec le langage de programmation Visual Basic 6. Visual Basic est l'un des langages de base au moyen desquels la plupart des logiciels de Microsoft sont produits. TIRAGE 2.1 est compilé et peut donc être utilisé sous le système d'exploitation Windows sans que l'on ait besoin de disposer du logiciel Visual Basic 6.

TIRAGE 2.1 utilise le gestionnaire de bases de données ACCESS pour loger les résultats des tirages ainsi que les sources des données à partir desquelles les tirages sont effectués. Les fichiers des sources des données et des résultats des tirages sont donc des tables du logiciel ACCESS. Cela constitue certainement une contrainte pour l'emploi de TIRAGE 2.1, car

¹ Après l'obtention d'une maîtrise de Mathématiques en 1965 à la Faculté des Sciences de Dijon, France, Julien Amegandjin a fait des études de Statistique au CESD et à l'ENSAE (1965-68) ainsi qu'à la Faculté des Sciences de Paris, Quai Saint Bernard (1967-70). Il est titulaire du diplôme d'Ingénieur Statisticien Economiste du CESD (Juin 1968), d'un Diplôme d'Etudes Approfondies de Statistique Mathématique (Juin 1967) et d'un Doctorat de Troisième Cycle de Statistique Mathématique (Juin 1970) de la Faculté des Sciences de Paris. Il a enseigné la Statistique pendant 17 années au CESD (1968-72), à l'ENSAE (1969-71), à l'Université de Paris-Dauphine (1968-72), à l'Université de Yaoundé (1973-75) et à l'IFORD (1973-85).

Il fut Directeur adjoint et Directeur des études du CESD (1970-72), puis Directeur de l'IFORD (1975-85). Il a eu une longue pratique des enquêtes par sondage, et particulièrement, dans le cadre des projets de développement des statistiques agricoles ou des projets de recensement agricole qu'il a dirigés pour le compte de la FAO de 1986 à 1999. Les fonctions qu'il a exercées de 1973 à 1999 le furent avec un statut de fonctionnaire des Nations Unies avec trois agences : le Département de la Coopération Technique du Secrétariat Général, DTCD (1973-79), la Commission Economique des Nations Unies pour l'Afrique, CEA (1980-85) et l'Organisation des Nations Unies pour l'Alimentation et l'Agriculture, FAO (1986-99).

Parmi les travaux ou publications de l'auteur, on peut citer :

- L'admissibilité des estimateurs et la statistique d'ordre dans la théorie de l'échantillonnage sur des populations finies (thèse de doctorat, Faculté des Sciences de Paris, Quai Saint Bernard, Juin 1970) ;
- Démographie Mathématique (manuel publié chez Economica dans la collection "Economie et Statistiques Appliquées" de l'Ecole Nationale de la Statistique et de l'Administration Economique, 265 pages, Septembre 1989).

L'auteur est à la retraite depuis l'année 2000 et réside maintenant au Togo, son pays natal. Il a conservé quelques activités et intervient occasionnellement comme consultant en Sondage. Il a notamment réalisé les plans de sondage de l'Enquête Agricole Annuelle de Côte d'Ivoire (Octobre 2003), de l'enquête QUIBB du Togo (Décembre 2003), du Recensement Général de l'Agriculture et du Cheptel du Niger (Avril 2004) et d'enquêtes MICS3 (Togo, Burundi et Mauritanie). Il a aidé à réviser le plan de sondage et à mettre au point les méthodes d'estimation pour la Deuxième Enquête sur la Consommation et le Secteur Informel du Tchad (ECOSIT2) en mai 2005. Il a eu également des interventions en 2005 et 2006 comme consultant en Sondage dans le cadre des enquêtes MICS3 de l'UNICEF.

Il a assuré des ateliers de formation en sondage au Mali (2007), au Congo (2008) et au Niger (2008). Il est engagé avec l'UNICEF (Bureaux régionaux de Dakar et d'Amman) pour donner des appuis à certains pays en faveur de l'élaboration des plans de sondage des enquêtes MICS de la 4^e édition (2009-2011).

Il est lauréat du Prix africain de Statistique décerné sous les auspices de la Commission Africaine des Statistiques des Nations Unies en janvier 2008.

l'utilisateur de ce logiciel doit savoir présenter ses sources de données sous forme de tables ACCESS. Il doit également être en mesure d'accéder aux résultats des tirages et d'en faire toutes les utilisations possibles.

Une connaissance minimum des tables ACCESS est donc exigée de l'utilisateur de TIRAGE 2.1. L'ouverture, la consultation et la fermeture d'une table doivent être maîtrisées au préalable par l'utilisateur de ce logiciel. La connaissance des notions de structure, de champ et de clé primaire d'une table est également indispensable ainsi que celle des notions de tri et de filtre des enregistrements. L'utilisateur de TIRAGE 2.1 doit enfin savoir comment importer la source des données d'un tirage dans la base de données Data.mdb associé au logiciel et comment exporter une table hors de cette base.

2. A propos de l'installation du logiciel

Pour l'utilisation du logiciel, il est nécessaire que soit créé un dossier TIRAGE sur le disque C, soit C:\TIRAGE, pour loger la base de données ACCESS liée au logiciel et nommée Data.mdb.

Au moment de l'installation du logiciel, le répertoire C:\Program Files\ est suggéré pour le recevoir. L'utilisateur pourra accepter ce répertoire ou en créer un autre. Il doit ensuite aller dans le répertoire C:\Program Files\Tirage\Data.mdb (s'il a accepté ce répertoire) pour réaliser une copie de la base Data.mdb qu'il ira coller dans C:\TIRAGE\.

3. Principe de la génération des nombres aléatoires et des tirages aléatoires

Avec le logiciel Visual Basic, un nombre aléatoire se génère au moyen de la fonction Rnd. L'exécution de la fonction Rnd fournit un nombre aléatoire suivant la loi uniforme sur l'intervalle [0,1]. A partir d'un tel nombre, on peut obtenir tout autre nombre aléatoire répondant à des conditions données. Quand on génère un ensemble de nombres aléatoires d'un même type, c'est-à-dire répondant aux mêmes conditions, on génère un échantillon aléatoire de nombres de ce type.

Soit U un univers composé de M unités et duquel on désire sélectionner au hasard un échantillon de m unités ou un échantillon de taille m. Les M unités statistiques de l'univers sont numérotées de 1 à M et ces numéros constituent les identités ou étiquettes des unités. Tirer un échantillon aléatoire de taille m de l'univers U consistera simplement à générer, à l'aide de la fonction Rnd de Visual Basic, m nombres compris entre 1 et M qui identifieront les unités sélectionnées. La génération des m nombres tiendra compte bien évidemment des spécificités du mode particulier de tirage considéré.

Ainsi, à l'aide du procédé de génération des nombres aléatoires, on peut répartir au hasard un jeu de cartes entre quatre joueurs. De nombreuses autres opérations aléatoires peuvent être mises en œuvre grâce au même procédé.

4. Les différents modes de tirage d'échantillons

Deux paramètres sont essentiels dans la définition du mode de tirage d'un échantillon : l'égalité ou non des probabilités de tirage attribuées aux unités de l'univers et la possibilité ou non de remise pour les unités tirées.

Un tirage avec probabilités égales attribue à l'avance une même probabilité à chaque unité statistique de l'univers sondé. Dans le cas où des probabilités inégales sont attribuées aux unités statistiques, on a affaire à un tirage avec probabilités inégales.

Les tirages ont lieu avec remise lorsque, à chaque tirage, l'unité statistique obtenue est remise dans l'univers avant le tirage suivant. C'est le cas des tirages où une même unité statistique peut figurer plus d'une fois dans l'échantillon.

Avec les tirages sans remise, une unité statistique tirée n'est pas remise dans l'univers avant de procéder aux tirages suivants. Dans ce cas, une unité statistique ne peut figurer plus d'une fois dans l'échantillon.

Deux modes de tirage d'échantillon très connus sont ceux associés au sondage aléatoire simple. Un sondage aléatoire simple (simple random sampling) est une méthode de tirage qui consiste à tirer dans une population finie de M unités, un échantillon de taille fixée m à partir des seuls identifiants des unités de telle façon que chaque unité ait la même probabilité d'appartenir à l'échantillon. Le sondage aléatoire simple peut être avec remise ou sans remise. Dans ce qui suit, nous parlerons de sondage aléatoire simple avec remise ou de sondage aléatoire simple sans remise. Et dans l'un ou l'autre cas, il s'agit de sondage à probabilités égales ou de tirage avec probabilités égales.

TIRAGE 2.1 utilise pour ces deux cas de tirage aléatoire, les symboles "tirage PEAR" et "tirage PESR". Le tirage PEAR² (à probabilités égales et avec remise) désigne le tirage aléatoire simple avec remise, ou ce qui est équivalent, le tirage avec probabilités égales et avec remise. Le tirage PESR³ (à probabilités égales et sans remise) désigne le tirage aléatoire simple sans remise ou encore le tirage avec probabilités égales et sans remise.

Dans toute la suite, le tirage PESR se rapporte exclusivement au sondage aléatoire simple sans remise et non pas au tirage systématique avec probabilités égales, un mode de tirage presque équivalent mais relevant d'un modèle probabiliste différent.

Le logiciel TIRAGE 2.1 utilise également les trois modes de tirage suivants :

- le tirage avec probabilités inégales et avec remise ou tirage PIAR⁴ ;
- le tirage systématique avec probabilités égales ou tirage SPE⁵ ;
- le tirage systématique avec probabilités inégales ou tirage SPI, et le cas particulier de ce dernier, qui est le tirage systématique avec probabilités proportionnelles à la taille des unités ou tirage SPPT.

Bien souvent, les probabilités inégales attribuées aux unités par un plan de sondage sont des probabilités proportionnelles à la taille des unités. Si les unités de l'univers sont, par exemple, des villages, le nombre de ménages du village peut servir de mesure de taille pour les villages.

Le tirage systématique avec probabilités égales (ou tirage SPE) est une variante du tirage aléatoire simple sans remise pour ce qui est de la constitution de l'échantillon bien que les deux méthodes de tirage ne soient pas totalement équivalentes parce que correspondant à des modèles probabilistes différents.

5. Les différents groupes de modules du logiciel

Trois groupes de modules (ou applications) composent le logiciel TIRAGE 2.1 :

- des modules relatifs à la génération de nombres aléatoires ;

² PEAR est mis pour "à Probabilités Egales et Avec Remise " selon Jean-Marie Grosbras

³ PESR est mis pour "à Probabilités Egales et Sans Remise " selon Jean-Marie Grosbras

⁴ PIAR est mis pour "à Probabilités Inégales et Avec Remise " selon Jean-Marie Grosbras

⁵ SPE, SPI et SPPT sont des symboles dus à J. Amegandjin

- des modules relatifs à des tirages d'échantillons au sein d'une population finie ;
- des modules utilitaires.

Les cinq modules suivants se rapportent à la génération de nombres aléatoires :

- Génération de nombres de l'intervalle $[0,1[$
- Génération selon le mode PEAR de nombres de 1 chiffre
- Génération selon le mode PEAR de nombres de K chiffres de type texte
- Génération selon le mode PEAR de nombres entiers compris entre M et K inclus
- Génération selon le mode PESR de nombres entiers compris entre M et K inclus

Les modules utilitaires qui ne sont pas des modules de génération de nombres aléatoires sont au nombre de six. Ce sont :

- Attribuer un numéro d'ordre aux unités d'une population
- Attribuer un numéro d'ordre aux US de chaque UP d'une base de sondage
- Compter les unités secondaires de chaque UP d'une base de sondage
- Cumuler les tailles des unités d'une population
- Réordonner au hasard les unités d'une population
- Calcul et contrôle des probabilités d'inclusion dans un tirage SPPT_{cpi}

Les modules se rapportant aux tirages d'échantillons aléatoires dans une population finie sont de deux types. Ils concernent, soit des tirages dans un sondage à un degré, soit des tirages dans un sondage à 2 degrés. Les modules relatifs aux tirages à un degré utilisés ici sont :

- le tirage selon le mode PEAR d'un échantillon d'unités
- le tirage selon le mode PESR d'un échantillon d'unités
- le tirage systématique avec probabilités égales d'un échantillon d'unités (tirage SPE)
- le tirage systématique avec probabilités proportionnelles à la taille des unités ou tirage SPPT
- le tirage systématique avec probabilités proportionnelles à la taille des unités basé sur le cumul des probabilités d'inclusion ou tirage SPPT_{cpi}

Trois modules relatifs aux tirages à deux degrés sont considérés dans TIRAGE 2.1 :

- un tirage selon le mode PESR au premier degré suivi d'un tirage selon le mode PESR d'un nombre constant d'unités secondaires dans chaque unité primaire ;
- un tirage selon le mode PIAR au premier degré suivi d'un tirage selon le mode PESR d'un nombre constant d'unités secondaires dans chaque unité primaire ;
- un tirage selon le mode SPPT au premier degré suivi d'un tirage selon le mode PESR d'un nombre constant d'unités secondaires dans chaque unité primaire.

6. La barre des menus du logiciel

Les différents modules de TIRAGE 2.1 présentés ci-dessus apparaissent au lancement du logiciel dans une barre de menus composés des éléments :

- ◆ Fichier
- ◆ Tirages avec remise
- ◆ Tirages sans remise
- ◆ Tirages dans un sondage à deux degrés
- ◆ Utilitaires

Le tableau "Barre des menus" ainsi que les images écran ci-après présentent les menus respectifs.

Barre des menus

Fichier

Quitter

Tirages avec remise

Génération PEAR d'un échantillon de nombres de 1 chiffre

Génération PEAR de nombres de K chiffres de type texte

Génération PEAR de nombres entiers compris entre M et Z inclus

Tirage PEAR d'un échantillon d'unités au sein d'une population

Tirage PIAR d'un échantillon d'unités au sein d'une population

Tirages sans remise

Génération PESR de nombres entiers compris entre M et Z inclus

Tirage PESR d'un échantillon d'unités au sein d'une population

Tirage systématique avec probabilités égales (SPE)

Tirage systématique avec probabilités proportionnelles à la taille des unités (SPPT)

Tirage systématique SPPT par cumul des probabilités d'inclusion (SPPTcpi)

Tirages dans un sondage à 2 degrés

Tirage PESR de l'échantillon du 1^{er} degré

Table source des tirages du 2^e degré à la suite de tirages PESR au 1^{er} degré

Tirage PESR de l'échantillon du 2^e degré avec un nombre constant d'US par UP

Tirage PIAR de l'échantillon du 1^{er} degré

Table source des tirages du 2^e degré à la suite de tirages PIAR au 1^{er} degré

Tirage PESR de l'échantillon du 2^e degré avec un nombre constant d'US par UP

Tirage SPPT de l'échantillon du 1^{er} degré

Table source des tirages du 2^e degré à la suite de tirages SPPT au 1^{er} degré

Tirage PESR de l'échantillon du 2^e degré avec un nombre constant d'US par UP

Utilitaires

Attribuer un numéro d'ordre aux unités d'une population

Attribuer un numéro d'ordre aux US de chaque UP d'une base de sondage

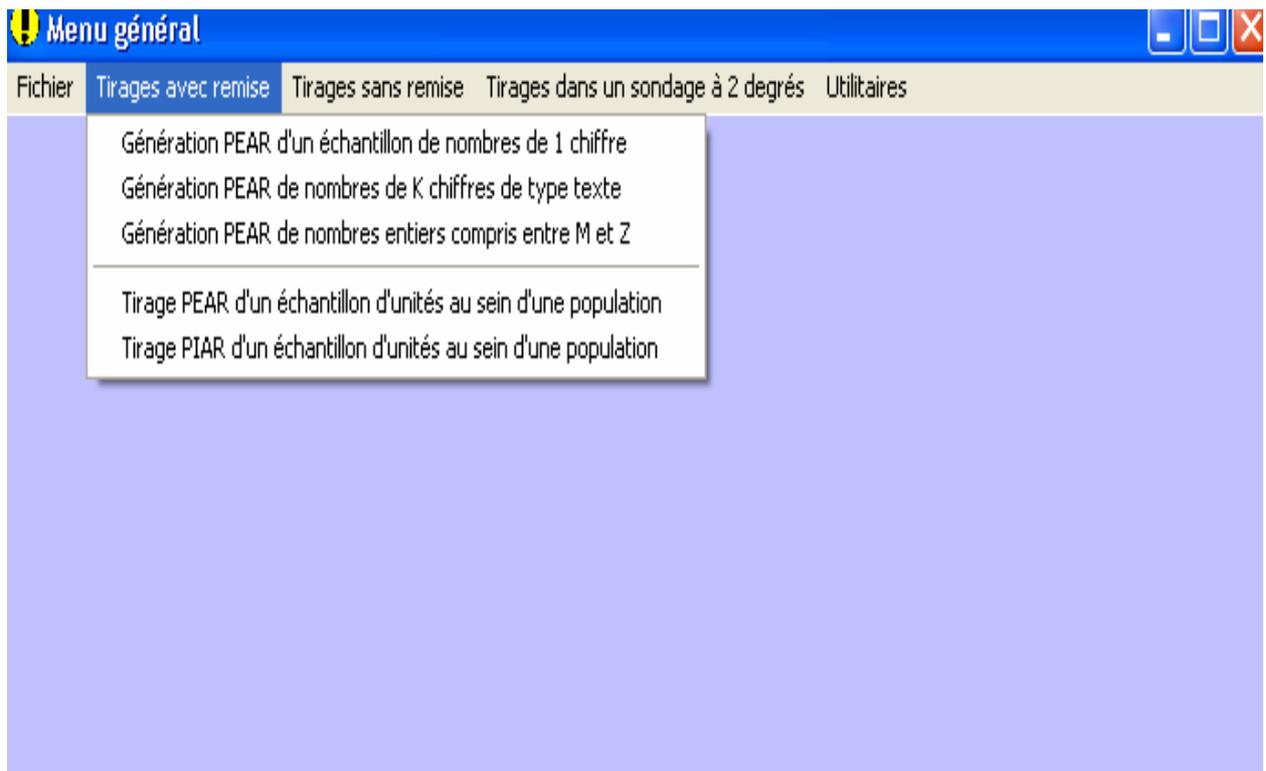
Compter les unités secondaires de chaque UP d'une base de sondage

Cumuler les tailles des unités d'une population

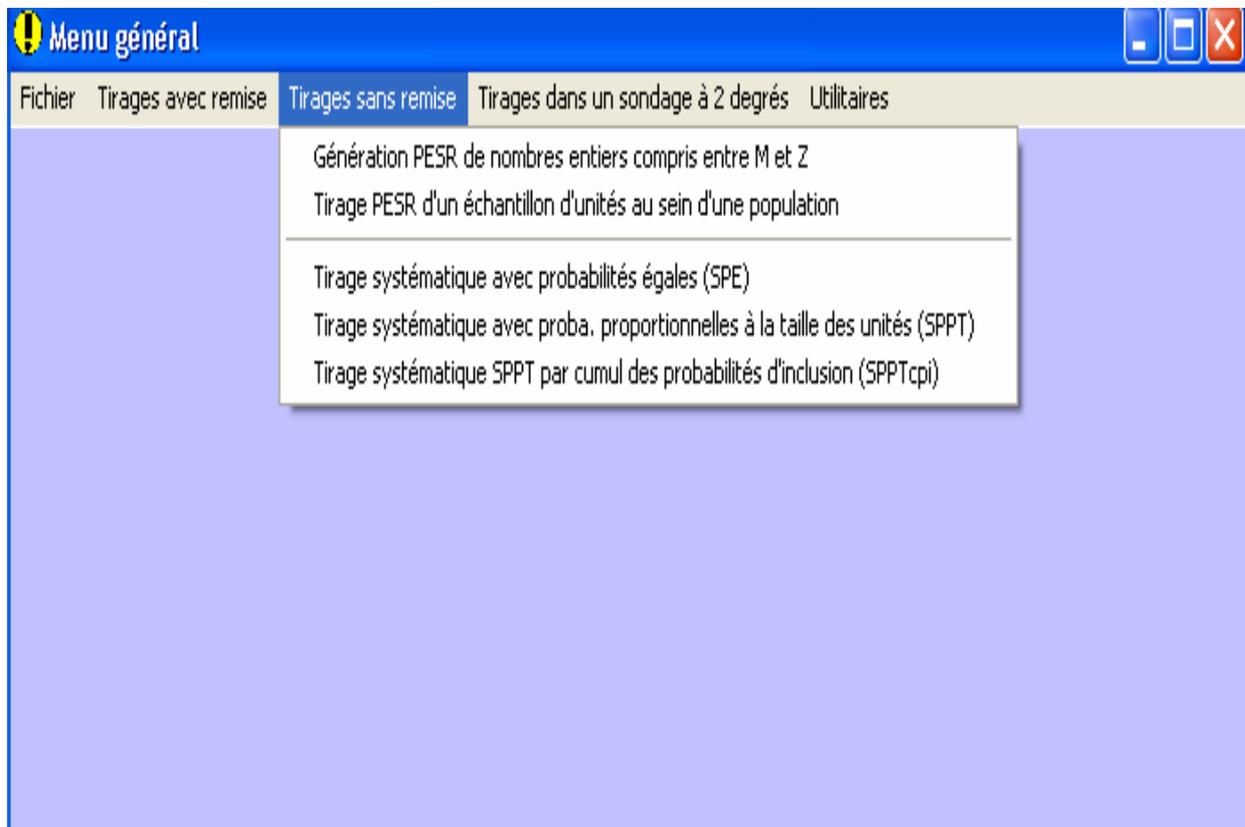
Réordonner au hasard les unités d'une population

Calcul et contrôle des probabilités d'inclusion dans un tirage SPPTcpi

Menu "Tirages avec remise"



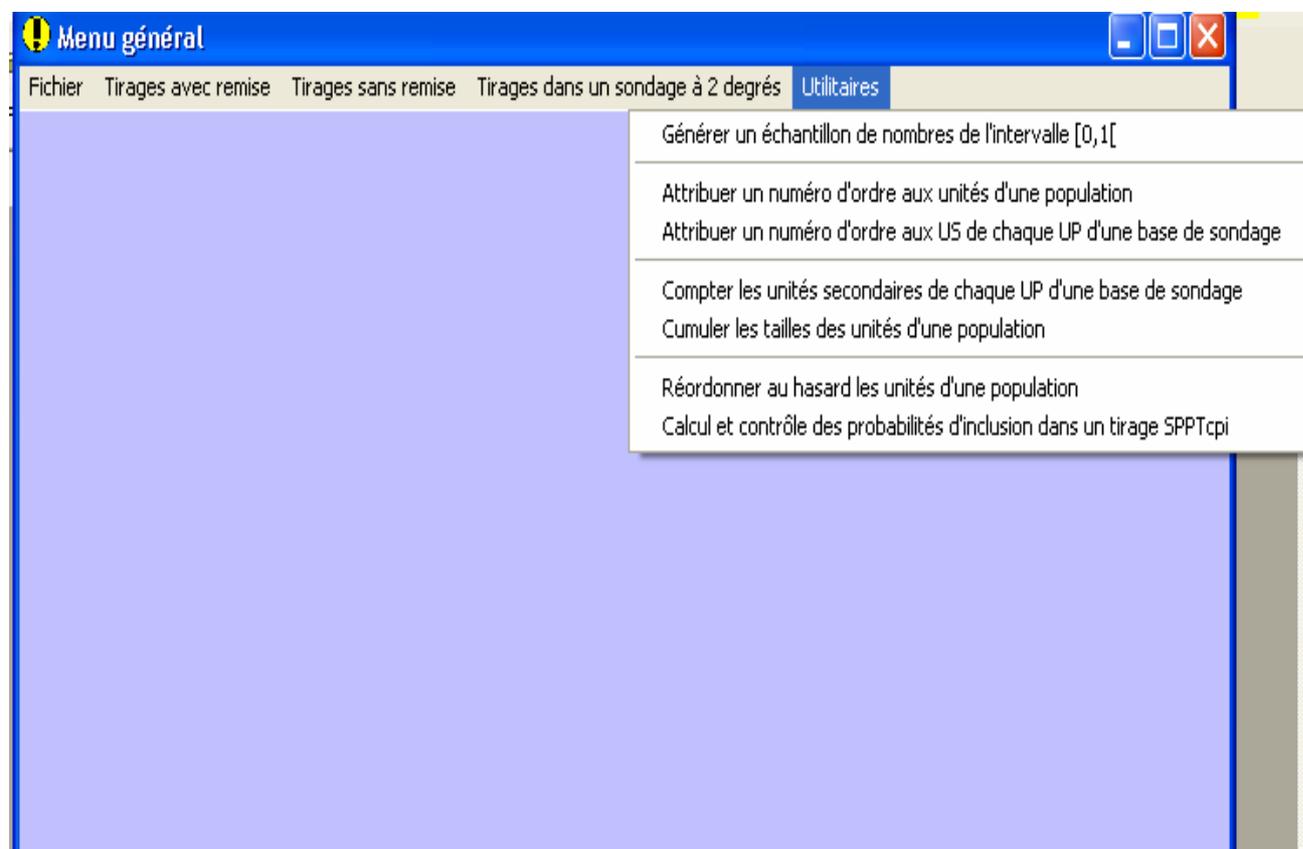
Menu "Tirages sans remise"



Menu "Tirages dans un sondage à 2 degrés"



Menu "Utilitaires"



Chapitre 2

Comment s'exécutent les modules du logiciel ?

Ce chapitre 2 est consacré au mode d'emploi du logiciel TIRAGE 2.1. Seulement quelques exemples d'application sont présentés ici au lieu d'une revue de la mise en œuvre de tous les modules.

Lorsque TIRAGE 2.1 est lancé, la page de la barre des menus s'ouvre et tout module d'un menu donné peut être exécuté. A l'exécution d'un module, une feuille Visual Basic se présente avec des zones de texte vierges à renseigner et un ou plusieurs boutons de commande destinés à exécuter des actions. Chaque fois, l'utilisateur doit renseigner les zones de texte en y inscrivant les informations recherchées pour l'exécution du module. Ensuite, il clique sur le bouton de commande pour exécuter l'action qui définit le module. Si plusieurs boutons de commande sont présents sur la feuille Visual Basic, l'opérateur clique sur ces boutons selon leur ordre d'arrivée sur la feuille du haut vers le bas.

Les exemples d'application présentés dans ce chapitre concernent :

- le module utilitaire "Cumuler les tailles des unités d'une population";
- le module "Génération PEAR d'un échantillon de nombres entiers compris entre M et Z inclus" ;
- le module "Tirage PIAR d'un échantillon d'unités au sein d'une population" ;
- des tirages d'un sondage à deux degrés faisant intervenir 3 modules liés : le module "Tirage systématique avec probabilités proportionnelles à la taille (tirage SPPT)" pour le 1^{er} degré, le module "Construire la table source des tirages du 2^e degré à la suite d'un tirage SPPT au 1^{er} degré" et le module "Tirage PESR de l'échantillon du 2^e degré avec un nombre constant d'US par UP" ;
- le module "Génération PEAR de nombres de K chiffres de type texte" ;
- la génération PESR d'un échantillon de nombres entiers de K chiffres ;
- le module utilitaire "Attribuer un numéro d'ordre aux unités d'une population";
- le module utilitaire "Réordonner au hasard les unités d'une population".

→ Remarque

Pour ces exemples d'application, nous utiliserons des bases de sondage qui sont des tables de la base de données ACCESS Data.mdb du répertoire C:\TIRAGE\. Ces tables possèdent toutes un champ Identité qui permet d'identifier les unités de la population représentée par la table. Le champ Identité est exclusivement de type numérique et non de type texte. Il en sera de même pour toutes les tables qui devront être importées dans la base de données Data.mdb pour être utilisées par TIRAGE 2.1. Si cette précaution n'est pas prise, le tri des enregistrements d'une table engendre des problèmes. Dans la base de données Data.mdb, le champ Identité est dénommé CODEZD pour les tables Dosso, Gaya, Madaoua et Magaria, et IDENTITE pour les tables ABCD, Maritime, Plateaux_Rural et Plateaux_Urbain.

Les tables de Data.mdb possèdent pour la plupart, un champ "Taille" indispensable pour les tirages qui doivent être effectués avec des probabilités proportionnelles à la taille des unités. Il devra en être de même pour les tables qui devront être importées dans la base de données Data.mdb à des fins de tirages avec des probabilités proportionnelles à la taille des unités. Le champ Taille s'appelle NBMEN pour les tables Dosso, Gaya, Madaoua et Magaria,

Nombre_Ménages pour la table ABCD et POPULATION pour les tables Maritime, Plateaux_Rural et Plateaux_Urbain.

Au champ "Taille" est souvent associé un champ Cumul qui contient les valeurs des cumuls des tailles des unités. Les modules de TIRAGE 2.1 intègrent la création du champ Cumul ainsi que le calcul des cumuls des tailles lorsque ce champ est requis.

1. Exemple 1 : Cumuler les tailles des unités d'une population

Une étape de certains tirages d'échantillons consiste à calculer le cumul des tailles des unités de la population sondée. C'est le cas du tirage PIAR (tirage avec probabilités inégales et avec remise) et du tirage systématique avec probabilités inégales, et en particulier, du tirage SPPT (tirage systématique avec probabilités proportionnelles à la taille des unités). Le module utilitaire "Cumuler les tailles des unités d'une population" permet de réaliser cette étape des tirages de cette catégorie.

Le module permet de créer en premier lieu un champ Cumul dans la table qui est la source des données pour le tirage à effectuer. Il exécute ensuite le calcul des cumuls des tailles des unités qui sont les valeurs du champ Cumul.

Feuille n° 1 : Cumuler les tailles des unités d'une population

Cumuler les tailles des unités statistiques

Navigation: [Left Arrow] [Add] [Right Arrow]

Base de données utilisée: C:\TIRAGE\Data.mdb

Nom de la table utilisée: Madaoua

Nom du champ "Identité de l'unité statistique": CODEZD

Nom du champ "Taille de l'unité statistique": NBMEN

Créer le champ CUMUL

Exécuter le calcul des cumuls

Au lancement du module, une feuille Visual Basic s'affiche avec 4 zones de texte vierges et 2 boutons de commande. Le nom de la base de données utilisée sera choisi par un double-clic sur le bouton de recherche situé à droite de la zone de texte correspondante et qui permet de parcourir les bases de données du disque dur. Ensuite, les autres informations seront affichées

dans les zones de texte respectives grâce à un clic sur les boutons de déroulement des listes situés à droite de ces zones de texte.

Créons le champ Cumul et calculons le cumul des tailles des unités pour la table Madaoua de la base de données Data.mdb. La table Madaoua est une base de sondage des zones de dénombrement (ZD) de Madaoua, un département du Niger en 2006.

L'opérateur double-clique sur le bouton de recherche situé à droite de la zone de texte en face de l'étiquette "Base de données utilisée". Il a alors accès au disque C dans lequel il sélectionne le répertoire TIRAGE, et dans ce répertoire, la base Data. Il clique ensuite sur le bouton de déroulement des listes de la zone de texte en face de l'étiquette "Nom de la table utilisée" et choisit Madaoua dans la liste déroulante qui s'affiche. Pour chacune des 2 zones de texte suivantes, il clique sur le bouton de déroulement des listes correspondant et il choisit les champs CODEZD et NBMEN respectivement pour l'identité et la taille de l'unité statistique. La feuille Visual Basic se présente alors sous la forme de la feuille n°1 ci-dessus.

Table Access n° 1 : Cumul des tailles des unités de la table Madaoua

	REGION	DEPT	CANTON	ZONE	CODEZD	NBMEN	CUMUL
	5	7	10	1	5710001	221	221
	5	7	10	2	5710002	243	464
	5	7	10	3	5710003	213	677
	5	7	10	4	5710004	191	868
	5	7	10	5	5710005	329	1197
	5	7	10	6	5710006	249	1446
	5	7	10	7	5710007	343	1789
	5	7	10	8	5710008	193	1982
	5	7	10	9	5710009	232	2214
	5	7	10	10	5710010	173	2387
	5	7	10	11	5710011	102	2489
	5	7	10	12	5710012	294	2783
	5	7	10	13	5710013	342	3125
	5	7	10	14	5710014	290	3415
	5	7	10	15	5710015	223	3638
	5	7	10	16	5710016	226	3864
	5	7	10	17	5710017	318	4182
▶	5	7	10	18	5710018	276	4458
	5	7	10	19	5710019	306	4764
	5	7	10	20	5710020	251	5015
	5	7	10	21	5710021	313	5328
	5	7	10	22	5710022	274	5602
	5	7	10	23	5710023	322	5924

L'opérateur clique ensuite sur le bouton de commande "Créer le champ Cumul" puis sur le bouton de commande "Exécuter le calcul des cumuls". L'exécution du module est terminée et l'opérateur peut en consulter le résultat en ouvrant la table Madaoua dans la base Data.mdb. Il doit constater que la table Madaoua comporte le nouveau champ créé avec les valeurs des cumuls des tailles comme le montre la table Access n°1 ci-dessus.

2. Exemple 2 : Génération selon le mode PEAR d'un échantillon de nombres entiers compris entre M et Z inclus

Ce module se trouve dans le menu "Tirage avec remise". Lorsqu'il est lancé, une feuille Visual Basic se présente avec 4 zones de texte vierges destinées à recevoir les informations concernant l'identité du tirage, le nombre limite inférieur M, le nombre limite supérieur Z et la taille de l'échantillon.

A titre d'exemple, les données suivantes ont été fournies par l'opérateur à la feuille Visual Basic :

- Identité du tirage : Toro
- Nombre limite inférieur inclus (M) : 17
- Nombre limite supérieur inclus (Z) : 229
- Taille de l'échantillon : 15

La feuille Visual Basic devient alors la feuille n°2 ci-dessous.

Feuille n° 2 : Génération PEAR d'un échantillon de nombres entiers compris entre M et Z inclus

Génération PEAR d'un échantillon de nombres compris entre M et Z inclus

Identité du tirage (de type Texte)	Toro
Nombre limite inférieur (M)	17
Nombre limite supérieur (Z)	229
Taille de l'échantillon	15

Exécuter le tirage

Nom de la table : TableENAMZ

L'opérateur clique ensuite sur le bouton de commande "Exécuter le tirage" et le module est exécuté. Le résultat du tirage est envoyé dans la table nommée TableENAMZ dont l'image écran est la table Access n°2 ci-dessous. Les résultats des 15 tirages successifs effectués constituent les enregistrements de la table TableENAMZ avec des champs donnant respectivement l'identité du tirage, le rang du tirage et le nombre aléatoire tiré.

Dans cet exemple, le nom de la table TableENAMZ est inscrit sur l'interface Visual Basic en bas et à droite. Il en est ainsi pour tous les modules : le nom de la table qui reçoit les résultats de l'exécution d'un module est toujours inscrit sur l'interface Visual Basic sauf si l'opération est une transformation interne d'une table comme c'est le cas de l'exécution du module "Cumuler les tailles des unités d'une population" présenté plus haut.

Notons que le champ donnant l'identité du tirage est fort utile, car la table TableENAMZ peut recevoir les résultats de différents tirages et ce champ permet de les distinguer.

Table Access n° 2 : Résultat du tirage (TableENAMZ)

Identité_Tirage	Rang_Tirage	Nombre_Tiré
Toro	1	49
Toro	2	116
Toro	3	162
Toro	4	137
Toro	5	106
Toro	6	121
Toro	7	190
Toro	8	32
Toro	9	184
Toro	10	71
Toro	11	148
Toro	12	50
Toro	13	100
Toro	14	188
Toro	15	165
	0	0

3. Exemple 3 : Tirage selon le mode PIAR d'un échantillon d'unités dans une population

Le module "Tirage PIAR d'un échantillon d'unités au sein d'une population" permet de tirer un échantillon d'unités au sein d'une population finie en attribuant des probabilités inégales aux unités, probabilités qui sont le plus souvent des probabilités proportionnelles aux tailles des unités comme ce sera le cas dans cet exemple 3.

La base de sondage qui va être utilisée dans cet exemple doit comporter un champ "Taille" pour contenir la taille des unités.

Le module du tirage PIAR exige la création d'un champ Cumul et le calcul des cumuls des tailles, lesquelles opérations sont intégrées au programme du module. L'opérateur n'a donc pas à exécuter au préalable le module "Cumuler les tailles des unités d'une population".

Au lancement du module, une feuille Visual Basic s'affiche avec 6 zones de texte vierges et 2 boutons de commande. L'opérateur doit fournir à cette feuille, les informations suivantes : nom de la base de données utilisée, nom de la table qui est la source des données ou base de sondage, nom du champ de la table source qui joue le rôle d'identité pour les unités statistiques, nom du champ de la table source qui joue le rôle de taille pour les unités statistiques, nombre de tirages à effectuer et identité du tirage.

Feuille n° 3 : Tirage PIAR d'un échantillon d'unités au sein d'une population

! Tirage PIAR d'un échantillon d'unités au sein d'une population

Base de données utilisée	C:\TIRAGE\Data.mdb
Table source des tirages	Dosso
Champ "Identité de l'unité statistique"	CODEZD
Champ "Taille de l'unité statistique"	NBMEN
Nombre de tirages à effectuer	25
Identité du tirage (Texte)	Dallas

Exécuter le tirage

Editer le résultat du tirage

Nom de la table : TablePIAR

Le nom de la base de données utilisée sera choisi au moyen d'un double-clic sur le bouton de recherche situé à droite de la zone de texte correspondante et qui permet de parcourir les bases de données du disque dur. Le nom de la table source ainsi que les noms des 2 champs de la table source seront inscrits dans les zones de texte correspondantes grâce à un clic sur les boutons de déroulement des listes situés à droite dans ces zones de texte. Le nombre de tirages à effectuer ainsi que l'identité du tirage seront ensuite inscrits dans les zones de texte correspondantes.

Pour cet exemple de tirage, utilisons la table Dosso de la base de données Data.mdb. La table Dosso est une base de sondage des ZD (zones de dénombrement) de Dosso, un département du Niger en 2004. Cette table remplit les conditions requises pour un tirage PIAR. Elle comporte bien un champ Identité nommé CODEZD et un champ Taille nommé NBMEN. Ces 2 caractéristiques de la table sont portées dans les zones de texte correspondantes. Ensuite, l'opérateur inscrit les deux informations suivantes sur la feuille Visual Basic :

Nombre de tirages à effectuer : 25
Identité du tirage : Dallas

Et la feuille Visual Basic prend la forme de la feuille n°3 ci-dessus. L'opérateur clique alors sur le bouton de commande "Exécuter le tirage". Le résultat du tirage est envoyé dans la table TablePIAR de la base Data.mdb dont l'image est la table Access n°3 ci-après.

Table Access n° 3 : Résultat du tirage PIAR (TablePIAR)

Identité_Tirage	Rang_Tirage	IDENTITE	REPRISE	TAILLE
Dallas	1	3110059	1	161
Dallas	2	3110252	1	162
Dallas	3	3110211	1	125
Dallas	4	3110003	1	111
Dallas	5	3110085	1	143
Dallas	6	3110263	1	122
Dallas	7	3110248	1	138
Dallas	8	3110223	1	110
Dallas	9	3110241	1	154
Dallas	10	3110172	1	158
Dallas	11	3110009	1	116
Dallas	12	3110031	1	109
Dallas	13	3110039	1	146
Dallas	14	3110248	2	138
Dallas	15	3110127	1	120
Dallas	16	3110220	1	129
Dallas	17	3190027	1	122
Dallas	18	3190028	1	93
Dallas	19	3110251	1	123
Dallas	20	3110161	1	145
Dallas	21	3110195	1	120
Dallas	22	3110005	1	137
Dallas	23	3190025	1	87
Dallas	24	3110089	1	138
Dallas	25	3110094	1	98

La table TablePIAR comporte deux champs particuliers : le champ Rang_Tirage et le champ Reprise. Les 25 tirages successifs obtenus sont tous identifiés par leur rang de tirage. Le champ "Reprise" indique le rang de tirage pour les unités qui sont tirées plus d'une fois. Ainsi, l'unité dont l'identité est 3110248 est tirée 2 fois, une première fois au tirage de rang n°7 et une deuxième fois, au de tirage de rang n°14. Le champ "Reprise" prend alors les valeurs 1 et 2 respectivement pour les 7^e et 14^e tirages.

En plus de la consultation du résultat du tirage dans la table TablePIAR, l'opérateur peut visualiser le résultat du tirage dans un état de Visual Basic comme le montre le tableau n°1 ci-après en cliquant sur le bouton de commande "Editer le résultat du tirage". Il peut imprimer cet état sans toutefois pouvoir en manipuler le fichier. Pour tout autre usage du résultat du tirage, l'opérateur devra retourner à la table TablePIAR dans la base Data.mdb.

Tableau n° 1 : Résultat du tirage PIAR selon un état Visual Basic

Identité du tirage	Rang du tirage	Identité_UP	Reprise	Taille_UP
Dallas	1	3110059	1	161
Dallas	2	3110252	1	162
Dallas	3	3110211	1	125
Dallas	4	3110003	1	111
Dallas	5	3110085	1	143
Dallas	6	3110263	1	122
Dallas	7	3110248	1	138
Dallas	8	3110223	1	110
Dallas	9	3110241	1	154
Dallas	10	3110172	1	158
Dallas	11	3110009	1	116
Dallas	12	3110031	1	109
Dallas	13	3110039	1	146
Dallas	14	3110248	2	138
Dallas	15	3110127	1	120
Dallas	16	3110220	1	129
Dallas	17	3190027	1	122
Dallas	18	3190028	1	93
Dallas	19	3110251	1	123
Dallas	20	3110161	1	145
Dallas	21	3110195	1	120
Dallas	22	3110005	1	137
Dallas	23	3190025	1	87
Dallas	24	3110089	1	138
Dallas	25	3110094	1	98

4. Exemple 4 : Tirages dans un sondage à deux degrés

Le présent exemple de tirages se rapporte à un sondage à deux degrés. Le logiciel TIRAGE 2.1 propose trois modes de sondage à deux degrés dont l'un est développé ici. Pour la réalisation des tirages d'un sondage à deux degrés, TIRAGE 2.1 suggère trois étapes : le tirage des unités primaires (ou du 1^{er} degré), la construction d'une table qui fournira la source des données pour le tirage du 2^e degré et le tirage des unités secondaires (ou du 2^e degré).

L'exemple traité ici concerne le troisième mode de sondage à 2 degrés du menu "Tirages dans un sondage à 2 degrés". Il consiste à tirer au 1^{er} degré, les unités selon le mode de tirage systématique avec probabilités proportionnelles à la taille des unités (tirage SPPT) et à tirer au 2^e degré, selon le mode PESR, un nombre constant d'unités secondaires par unité primaire. Mais avant le tirage du 2^e degré, un module du logiciel permet de construire une table dénommée SourceTirages2eD_SPPT où sont rassemblées les données qui définissent les paramètres du tirage du 2^e degré. Elles comportent aussi bien des résultats du tirage du 1^{er} degré que certains choix du sondeur.

Nous allons donc utiliser successivement les trois modules suivants :

- Tirage SPPT de l'échantillon du 1^{er} degré ;
- Table source des tirages du 2^e degré à la suite de tirages SPPT au 1^{er} degré ;
- Tirage PESR de l'échantillon du 2^e degré avec un nombre constant d'US par UP.

Etape 1

Tirage systématique avec probabilités proportionnelles à la taille des unités (tirage SPPT)

Comme pour le tirage PIAR, le tirage SPPT requiert de la base de sondage de comporter un champ "Taille". Le tirage SPPT requiert, en outre, la création d'un champ Cumul ainsi que le calcul des cumuls des tailles. Les deux opérations sont exécutées directement par le module et l'opérateur n'a pas à utiliser au préalable le module "Cumuler les tailles des unités d'une population" à cette fin.

Pour cet exemple de tirage, nous allons utiliser la table Madaoua de la base de données Data.mdb.

Au lancement du module, une feuille Visual Basic s'affiche avec 6 zones de texte vierges et 2 boutons de commande. L'opérateur doit fournir à cette feuille, les informations suivantes : le nom de la base de données utilisée, le nom de la table qui sert de base de sondage, le nom du champ de la table qui joue le rôle d'identité pour les unités statistiques, le nom du champ de la table qui joue le rôle de taille pour les unités statistiques, le nombre de tirages à effectuer et l'identité du tirage.

Le nom de la base de données utilisée sera choisi au moyen d'un double-clic sur le bouton de recherche situé à droite de la zone de texte correspondante et qui permet de parcourir les bases de données du disque dur. Le nom de la table source ainsi que les noms des 2 champs de la table source seront inscrits dans les zones de texte correspondantes grâce à un clic sur les boutons de déroulement des listes situés à droite dans ces zones de texte. Le nombre de tirages à effectuer ainsi que l'identité du tirage seront ensuite inscrits dans les zones de texte correspondantes.

Après avoir sélectionné la base de données Data, la table Madaoua et les données requises de cette table, l'opérateur inscrit les deux informations suivantes sur la feuille Visual Basic :

Nombre de tirages à effectuer : 20
Identité du tirage : Golfe

La feuille Visual Basic se présente sous la forme de la feuille n°4 ci-dessous. Ensuite, l'opérateur clique sur le bouton de commande "Exécuter le tirage" et le module est exécuté. Le résultat du tirage est envoyé dans la table TableSPPT de la base Data.mdb dont l'image est la table Access n°4 ci-dessous.

Feuille n° 4 : Tirage SPPT d'un échantillon d'unités au sein d'une population

The screenshot shows a Visual Basic form with a blue title bar containing a yellow warning icon and the text "Tirage SPPT d'un échantillon d'unités". The form has a light beige background and contains several input fields and buttons. The fields are arranged in a vertical list on the left, with their corresponding values in text boxes or dropdown menus on the right. At the bottom, there are three buttons: "Exécuter le tirage", "Editer le résultat du tirage", and a label "Nom de la table : TableSPPT".

Base de données utilisée	C:\TIRAGE\Data.mdb
Table source des tirages	Madaoua
Champ "Identité de l'unité statistique"	CODEZD
Champ "Taille de l'unité statistique"	NBMEN
Taille de l'échantillon à tirer	20
Identité du tirage (Texte)	Golfe

Exécuter le tirage

Editer le résultat du tirage

Nom de la table : TableSPPT

Table Access n° 4 : Résultat du tirage SPPT (TableSPPT)

Identité_Tirage	Rang_Tirage	Nombre_Tiré	Identité_Unité	Taille_Unité	Cumul_Modifié
Golfe	1	7357	5710002	243	9280
Golfe	2	48020	5710011	102	49780
Golfe	3	88683	5710018	276	89160
Golfe	4	129346	5710025	460	131980
Golfe	5	170009	5710035	263	170540
Golfe	6	210672	5710043	283	213240
Golfe	7	251335	5710050	316	251760
Golfe	8	291998	5710060	224	294160
Golfe	9	332661	5710069	244	337200
Golfe	10	373324	5710078	214	375100
Golfe	11	413987	5710086	399	414200
Golfe	12	454650	5710095	303	460020
Golfe	13	495313	5710103	294	497200
Golfe	14	535976	5710111	336	539380
Golfe	15	576639	5710121	213	580100
Golfe	16	617302	5710130	181	617700
Golfe	17	657965	5710140	271	660500
Golfe	18	698628	5710151	248	703360
Golfe	19	739291	5710164	230	743440
Golfe	20	779954	5710177	208	782800
		0			

En plus de la consultation du résultat du tirage dans une table Access (TableSPPT), l'opérateur peut visualiser le résultat du tirage dans un état de Visual Basic en cliquant sur le bouton de commande "Editer le résultat du tirage". Il obtient le tableau n°2 ci-après. Il peut imprimer cet état sans toutefois pouvoir en manipuler le fichier. Pour tout autre usage du résultat du tirage, l'opérateur devra retourner à la table TableSPPT dans la base Data.mdb.

Tableau n° 2 : Résultat du tirage SPPT selon un état Visual Basic

Echantillon tiré				
Identité_Tirage	Rang_Tirage	Nombre_Tiré	Identité_Unité	Taille_Unité
Golfe	1	7357	5710002	243
Golfe	2	48020	5710011	102
Golfe	3	88683	5710018	276
Golfe	4	129346	5710025	460
Golfe	5	170009	5710035	263
Golfe	6	210672	5710043	283
Golfe	7	251335	5710050	316
Golfe	8	291998	5710060	224
Golfe	9	332661	5710069	244
Golfe	10	373324	5710078	214
Golfe	11	413987	5710086	399
Golfe	12	454650	5710095	303
Golfe	13	495313	5710103	294
Golfe	14	535976	5710111	336
Golfe	15	576639	5710121	213
Golfe	16	617302	5710130	181
Golfe	17	657965	5710140	271
Golfe	18	698628	5710151	248
Golfe	19	739291	5710164	230
Golfe	20	779954	5710177	208

Etape 2

Construction de la table source des tirages du 2^e degré à la suite d'un tirage SPPT au 1^{er} degré

A partir de données de la table TableSPPT et d'autres caractéristiques du tirage du 2^e degré qu'il se choisit, l'opérateur construit une table source des tirages du 2^e degré dénommée "SourceTirages2eD_SPPT" à l'aide du module "Table-source des tirages du 2^e degré à la suite de tirages SPPT au 1^{er} degré".

Au lancement de la construction de la table source des tirages du 2^e degré, une feuille Visual Basic s'affiche avec 2 zones de texte vierges destinées à recevoir respectivement l'identité du tirage du 1^{er} degré et le nombre constant d'unités secondaires à tirer par unité primaire. L'identité du tirage SPPT effectué au 1^{er} degré qui est Golfe sera affichée dans la première zone de texte. Le nombre constant d'unités secondaires à tirer par unité primaire sera inscrit dans la 2^e zone de texte. Choisissons ici ce nombre égal à 10. La feuille Visual Basic devient alors la feuille n°5 ci-dessous. La construction de la table est ensuite exécutée par un clic sur le bouton de commande "Exécuter la construction de la table". Le résultat de la construction de la table est la table SourceTirages2eD_SPPT figurée par la table Access n°5 ci-dessous.

Feuille n° 5 : Construction de la table source des tirages du 2^e degré à la suite d'un tirage SPPT au 1^{er} degré

Table Access n° 5 : Table SourceTirages2eD_SPPT

Identité_UP	Taille_UP	NombreUS_àTirer
5710002	243	10
5710011	102	10
5710018	276	10
5710025	460	10
5710035	263	10
5710043	283	10
5710050	316	10
5710060	224	10
5710069	244	10
5710078	214	10
5710086	399	10
5710095	303	10
5710103	294	10
5710111	336	10
5710121	213	10
5710130	181	10
5710140	271	10
5710151	248	10
5710164	230	10
5710177	208	10
		0

Etape 3

Tirage PESR de l'échantillon du 2^e degré

Après la construction de la table SourceTirages2eD_SPPT, l'opérateur procède au tirage du 2^e degré à l'aide du module " Tirage PESR de l'échantillon du 2^e degré avec un nombre constant d'US par UP". Au lancement du module, une feuille Visual Basic se présente avec 6 zones de texte vierges et 2 boutons de commande. L'opérateur devra fournir à la feuille le nom de la base de données utilisée, le nom de la table qui est la source des tirages, les noms des champs de la table source qui jouent respectivement le rôle d'identité et le rôle de taille pour l'unité primaire, le nom du champ de la table source qui contient le nombre de tirages à effectuer par unité primaire et l'identité du tirage à exécuter.

Feuille n° 6 : Tirage PESR d'un échantillon du 2^e degré à la suite d'un tirage SPPT

! Tirage PESR d'un échantillon du 2e degré à la suite d'un tirage SPPT

Base de données utilisée	C:\TIRAGE\Data.mdb
Table source des tirages	SourceTirages2eD_SPPT
Champ "Identité de l'unité primaire"	Identité_UP
Champ "Taille de l'unité primaire."	Taille_UP
Champ "Nombre d'US à tirer"	NombreUS_àTirer
Identité du tirage (de type Texte)	Lotus

Exécuter le tirage

Editer le résultat du tirage

Nom de la table : RésultatTirages2eD_SPPT

L'opérateur choisira la base Data dans le répertoire TIRAGE du disque C en double-cliquant sur le bouton de recherche situé à droite de la zone de texte correspondante et qui permet de parcourir les bases de données du disque dur. Puis, il choisira la table SourceTirages2eD_SPPT dans la liste déroulante qui s'affiche quand il aura cliqué sur le bouton de déroulement des listes situé à droite dans la zone de texte correspondante. Ensuite, l'opérateur fait afficher successivement les listes déroulantes des 3 zones de texte suivantes pour sélectionner les noms des 3 champs de la table SourceTirages2eD_SPPT recherchés. Enfin, il inscrit dans la zone de texte concernée, le nom de l'identité du tirage à exécuter, soit Lotus. On obtient alors les données du tirage du 2^e degré comme l'indique la feuille n°6 ci-dessus.

Ensuite, l'opérateur clique sur le bouton de commande "Exécuter le tirage". Le résultat du tirage est envoyé dans la table RésultatTirages2eD_SPPT dont une image écran partielle est présentée par la table Access n°6 ci-dessous. En plus de cette table, un état de Visual Basic obtenu par un clic sur le bouton de commande "Editer le résultat du tirage", permet de visualiser le résultat du tirage. Le tableau n°3 ci-dessous donne une image écran partielle de cet état Visual Basic.

Table Access n° 6 : Table RésultatTirages2eD_SPPT

IDTirage	Rang_Tirage	NuméroOrdre_Tiré	Identité_UP	Taille_UP
Lotus	10	3	5710002	243
Lotus	8	28	5710002	243
Lotus	2	29	5710002	243
Lotus	5	35	5710002	243
Lotus	7	48	5710002	243
Lotus	9	61	5710002	243
Lotus	4	99	5710002	243
Lotus	3	115	5710002	243
Lotus	1	160	5710002	243
Lotus	6	186	5710002	243
Lotus	3	2	5710011	102
Lotus	1	4	5710011	102
Lotus	5	8	5710011	102
Lotus	6	53	5710011	102
Lotus	4	54	5710011	102
Lotus	10	65	5710011	102
Lotus	2	77	5710011	102
Lotus	9	81	5710011	102
Lotus	7	83	5710011	102
Lotus	8	88	5710011	102
Lotus	7	18	5710018	276
Lotus	10	24	5710018	276
Lotus	4	43	5710018	276

Enr : 22 sur 200

Tableau n° 3 : Résultat du tirage du 2^e degré selon un état Visual Basic

Identité_Tirage	Rang_Tirage	Numéro d'ordre tiré	Identité_UP	Taille_UP
Lotus	10	3	5710002	243
Lotus	8	28	5710002	243
Lotus	2	29	5710002	243
Lotus	5	35	5710002	243
Lotus	7	48	5710002	243
Lotus	9	61	5710002	243
Lotus	4	99	5710002	243
Lotus	3	115	5710002	243
Lotus	1	160	5710002	243
Lotus	6	186	5710002	243
Lotus	3	2	5710011	102
Lotus	1	4	5710011	102
Lotus	5	8	5710011	102
Lotus	6	53	5710011	102
Lotus	4	54	5710011	102
Lotus	10	65	5710011	102
Lotus	2	77	5710011	102
Lotus	9	81	5710011	102
Lotus	7	83	5710011	102
Lotus	8	88	5710011	102
Lotus	7	18	5710018	276

On observe que dans la table RésultatTirages2eD_SPPT ou dans l'état Visual Basic précédent, les 10 unités secondaires tirées par unité primaire sont triées selon l'ordre croissant du champ NuméroOrdre_Tiré et non selon le champ Rang_Tirage. La raison est simplement de pouvoir disposer les unités secondaires tirées selon leur proximité sur le terrain.

5. Exemple 5 : Génération PEAR de nombres de K chiffres de type texte

Le module "Génération PEAR de nombres de K chiffres de type texte" appartient au menu "Tirages avec remise". Il permet d'obtenir au hasard des nombres de K chiffres semblables aux nombres d'une table de nombres au hasard commençant éventuellement par un chiffre 0 ou plusieurs. C'est ce type de nombre qui est appelé ici un nombre de type texte (ou alphanumérique). L'utilité du module est de générer, en cas de besoin, une table de nombres au hasard de K chiffres dont le fichier peut être reproduit facilement pour un usage sur le terrain.

Pour générer un nombre de K chiffres de type texte, on générera un nombre compris entre 0 et 1 selon la loi uniforme sur $[0,1[$. Ce nombre sera multiplié par 10^K , et on retiendra la partie entière du résultat de la multiplication. L'on prendra le soin d'ajouter le chiffre 0 autant de fois que cela est nécessaire devant la partie entière obtenue pour avoir K chiffres à la suite.

A titre d'exemple, générons un échantillon PEAR de taille n de nombres de K chiffres de type texte. Nous allons appliquer n fois la méthode de génération d'un nombre de K chiffres de type texte sans restriction, c'est-à-dire avec une possibilité de remise.

Feuille n° 7 : Génération PEAR d'un échantillon de nombres de K chiffres de type texte

The screenshot shows a software interface for generating a PEAR sample of K-digit text numbers. The window title is "Génération PEAR d'un échantillon de nombres de K chiffres de type texte". At the top, there is a navigation bar with a yellow warning icon, a "K" icon, left and right arrows, the text "Ado", and more left and right arrows. Below this are three input fields:

- "Identité du tirage (Texte)" with the value "Marine".
- "Longueur K des nombres à tirer" with the value "3".
- "Taille de l'échantillon à tirer" with the value "25".

Below these fields is a large button labeled "Exécuter le tirage". At the bottom right, there is a label "Nom de la table : TableENAK".

Au lancement du module, une feuille Visual Basic s'affiche avec 3 zones de texte vierges destinées à recevoir le nom de l'identité du tirage, la longueur K des nombres à tirer et la taille de l'échantillon à tirer. L'opérateur renseigne les zones de texte avec les données ci-après :

- Identité du tirage : Marine
- Longueur K des nombres à tirer : 3
- Taille de l'échantillon à tirer : 25

La feuille Visual Basic devient alors la feuille n°7 ci-dessus. Ensuite, l'opérateur clique sur le bouton de commande "Exécuter le tirage" et peut consulter le résultat du tirage dans la table TableENAK de la base de données Data.mdb. Une image écran partielle de cette table est la table Access n°7 ci-dessous.

Table Access n° 7 : TableENAK

	Identité_Tirage	Rang_Tirage	Nombre_Tiré
	Marine	1	667
	Marine	2	843
	Marine	3	002
	Marine	4	154
	Marine	5	748
	Marine	6	001
	Marine	7	577
	Marine	8	534
	Marine	9	543
	Marine	10	067
	Marine	11	273
	Marine	12	010
	Marine	13	676
	Marine	14	234
	Marine	15	330
	Marine	16	420
	Marine	17	948
	Marine	18	159
	Marine	19	582
	Marine	20	830
	Marine	21	476
	Marine	22	552
	Marine	23	717
	Marine	24	734
	Marine	25	088
*			0

Enr : 18 sur 25

6. Exemple 6 : Génération PESR d'un échantillon de nombres entiers de K chiffres

Pour générer un échantillon de nombres entiers distincts de K chiffres, on se servira du module "Génération PESR de nombres compris entiers entre M et Z inclus" avec les conditions :

$$M = 10^{K-1}$$

et

$$Z = 10^K - 1$$

A titre d'exemple, générons un échantillon PESR de 25 nombres entiers de 3 chiffres. Rappelons que le mode de tirage PESR est le tirage aléatoire simple sans remise qui consiste à effectuer des tirages avec probabilités égales et sans remise.

Au lancement du module "Génération PESR de nombres entiers compris entre M et Z inclus", une feuille Visual Basic s'affiche avec 4 zones de texte vierges destinées à recevoir le nom de l'identité du tirage, le nombre limite inférieur M, le nombre limite supérieur Z et la taille de l'échantillon à tirer. L'opérateur renseigne les zones de texte avec les données ci-après :

- Identité du tirage : Domino
- Nombre limite inférieur M : 100
- Nombre limite supérieur Z : 999
- Taille de l'échantillon à tirer : 25

Feuille n° 8 : Génération PESR d'un échantillon de nombres compris entre M et Z inclus

The screenshot shows a Visual Basic form with a blue title bar containing a warning icon and the text "Tirer selon le mode PESR un échantillon de nombres compris entre M et Z inclus". At the top center, there is a navigation bar with four buttons: a double left arrow, a single left arrow, the text "Adodd", a single right arrow, and a double right arrow. Below this are four rows of labels and text boxes:

Identité du tirage (de type Texte)	Domino
Nombre limite inférieur (M)	100
Nombre limite supérieur (Z)	999
Taille de l'échantillon	25

At the bottom center, there is a button labeled "Exécuter le tirage". In the bottom right corner, there is a label "Nom de la table : TablePesrENA".

La feuille Visual Basic devient alors la feuille n°8 ci-dessus. Ensuite, l'opérateur clique sur le bouton de commande "Exécuter le tirage" et peut consulter le résultat du tirage dans la table TablePesrENA de la base de données Data.mdb. Une image écran de cette table est la table Access n°8 ci-dessous.

Table Access n° 8 : TablePesrENA

	Identité_Tirage	Rang_Tirage	Nombre_Tiré
	Domino	1	528
	Domino	2	955
	Domino	3	642
	Domino	4	118
	Domino	5	769
	Domino	6	515
	Domino	7	370
	Domino	8	957
	Domino	9	429
	Domino	10	256
	Domino	11	231
	Domino	12	372
	Domino	13	441
	Domino	14	696
▶	Domino	15	736
	Domino	16	639
	Domino	17	543
	Domino	18	932
	Domino	19	650
	Domino	20	280
	Domino	21	731
	Domino	22	290
	Domino	23	829
	Domino	24	963
	Domino	25	206
*		0	0

Enr : 15 sur 25

7. Exemple 7 : Attribuer un numéro d'ordre aux unités d'une population

Il peut arriver qu'on veuille donner un numéro d'ordre aux unités d'une population et disposer d'une deuxième identité de ces unités en plus de leur identité de définition. Le module "Attribuer un numéro d'ordre aux unités d'une population" permet de réaliser la numérotation des unités de la population. Ce module appartient au menu "Utilitaires".

Pour exécuter le module, utilisons la table Gaya de la base de données Data.mdb. La table Gaya est une base de sondage des ZD (zones de dénombrement) du département de Gaya au Niger en 2004.

Au lancement du module, une feuille Visual Basic se présente avec 3 zones de texte vides et 2 boutons de commande. L'opérateur devra fournir à la feuille le nom de la base de données utilisée, le nom de la table utilisée et le nom du champ de la table qui joue rôle d'identité pour les unités.

Feuille n°9 : Attribuer un numéro d'ordre aux unités d'une population

! Attribuer un numéro d'ordre aux unités d'une population

Navigation: << < Add > >>

Nom de la base de données utilisée: C:\TIRAGE\Data.mdb

Nom de la table utilisée: Gaya

Nom du champ "identité de l'unité statistique": CODEZD

Créer le champ ORDRE

Créer les numéros d'ordre

En se servant du bouton de recherche des bases de données et du bouton de déroulement des listes, l'opérateur fournit les informations suivantes à la feuille Visual Basic :

- Nom de la base de données utilisée : Data
- Nom de la table utilisée : Gaya
- Nom du champ "Identité de l'unité statistique" : CODEZD

La feuille Visual Basic devient alors la feuille n°9 ci-dessus. Ensuite, l'opérateur clique sur le bouton de commande "Créer le champ ORDRE", puis sur le bouton de commande "Créer les numéros d'ordre". Le nouveau champ Ordre et ses valeurs apparaissent dans la table Gaya. Une image écran partielle de cette table est la table Access n°9 ci-dessous avec les unités de numéros d'ordre 51 à 73.

Table Access n° 9 : Table Gaya

	REGION	DEPT	CANTON	ZONE	CODEZD	NBMEN	ORDRE
	3	4	10	55	3410055	68	51
	3	4	10	56	3410056	148	52
	3	4	10	57	3410057	230	53
	3	4	10	58	3410058	227	54
	3	4	10	59	3410059	158	55
	3	4	10	60	3410060	89	56
	3	4	10	61	3410061	154	57
	3	4	11	1	3411001	110	58
	3	4	11	2	3411002	157	59
▶	3	4	11	3	3411003	143	60
	3	4	11	4	3411004	126	61
	3	4	11	5	3411005	148	62
	3	4	11	6	3411006	212	63
	3	4	11	7	3411007	129	64
	3	4	11	8	3411008	180	65
	3	4	11	9	3411009	84	66
	3	4	11	10	3411010	136	67
	3	4	11	11	3411011	288	68
	3	4	11	12	3411012	239	69
	3	4	11	13	3411013	219	70
	3	4	11	14	3411014	154	71
	3	4	11	15	3411015	182	72
	3	4	11	16	3411016	130	73

Enr : 60 sur 198

8. Exemple 8 : Réordonner au hasard les unités d'une population

Le module "Réordonner au hasard les unités d'une population" est un module utilitaire qui permet d'ordonner de façon tout à fait aléatoire les unités d'une base de sondage comme l'exigent certains modes de tirage préalablement aux tirages. C'est le cas de la méthode de tirage systématique de Hartley et Rao, un cas particulier de tirage systématique avec probabilités proportionnelles à la taille des unités (tirage SPPT).

Nous allons utiliser ce module pour réordonner les villages de la table ABCD de la base de données Data.mdb. La table ABCD est une base de sondage constituée de 20 villages imaginaires.

Au lancement du module, une feuille Visual Basic se présente avec 5 zones de texte vierges et un bouton de commande. L'opérateur devra fournir à la feuille le nom de la base de données utilisée, le nom de la table utilisée, les noms des champs de la table qui jouent le rôle d'identité et le rôle de taille pour les unités et l'identité du tirage à exécuter.

En se servant des boutons de recherche des bases de données ou de déroulement des listes, l'opérateur fournit les informations suivantes à la feuille Visual Basic :

Feuille n° 10 : Réordonner au hasard les unités d'une population

Réordonner au hasard les unités d'une population

Base de données utilisée: C:\TIRAGE\Data.mdb

Table source des tirages: ABCD

Champ "Identité de l'unité statistique": Identité

Champ "Taille de l'unité statistique": Nombre_Ménages

Identité du tirage (Texte): Palme

Exécuter le tirage

Nom de la table : TableOrdre

- Base de données utilisée : Data
- Table source des tirages : ABCD
- Champ "Identité de l'unité statistique" : Identité
- Champ "Taille de l'unité statistique" : Nombre_Ménages
- Identité du tirage (type texte) : Palme

La feuille Visual Basic devient alors la feuille n° 10 ci-dessus. Ensuite, l'opérateur clique sur le bouton de commande "Exécuter le tirage". Le résultat du tirage est envoyé dans la table TableOrdre de Data.mdb. Le champ Nouveau_Numéro de cette table fournit le nouvel ordre recherché pour les 20 villages. Une image écran de cette table est la table Access n° 10 ci-dessous.

Signalons que le module "Réordonner au hasard les unités d'une population" utilise le tirage PESR. Dans l'exemple traité ici, la méthode a consisté à tirer avec probabilités égales et sans remise, 20 nombres compris entre 1 et 20 inclus, ce qui est une façon de réaliser au hasard une permutation de l'ensemble {1, 2,, 20}. Le champ Nombre_Tiré fournit les 20 nombres tirés dans l'ordre de tirage, c'est-à-dire selon le rang de tirage qui n'est autre que le champ Nouveau_Numéro.

Table Access n° 10 : TableOrdre

	Identité_Tirage	Nombre_Tiré	Nouveau_Numéro	Identité_Unité	Taille_Unité
	Palme	19	1	8139	148
	Palme	15	2	8135	266
	Palme	11	3	8131	210
	Palme	12	4	8132	143
	Palme	18	5	8138	203
	Palme	7	6	8127	235
	Palme	20	7	8140	149
	Palme	17	8	8137	126
	Palme	2	9	8122	240
	Palme	4	10	8124	127
	Palme	6	11	8126	184
	Palme	8	12	8128	277
	Palme	13	13	8133	223
	Palme	9	14	8129	211
	Palme	10	15	8130	112
	Palme	3	16	8123	115
	Palme	5	17	8125	181
	Palme	16	18	8136	102
	Palme	1	19	8121	136
	Palme	14	20	8134	252
*		0	0	0	0

Enr : 14 sur 20

Chapitre 3

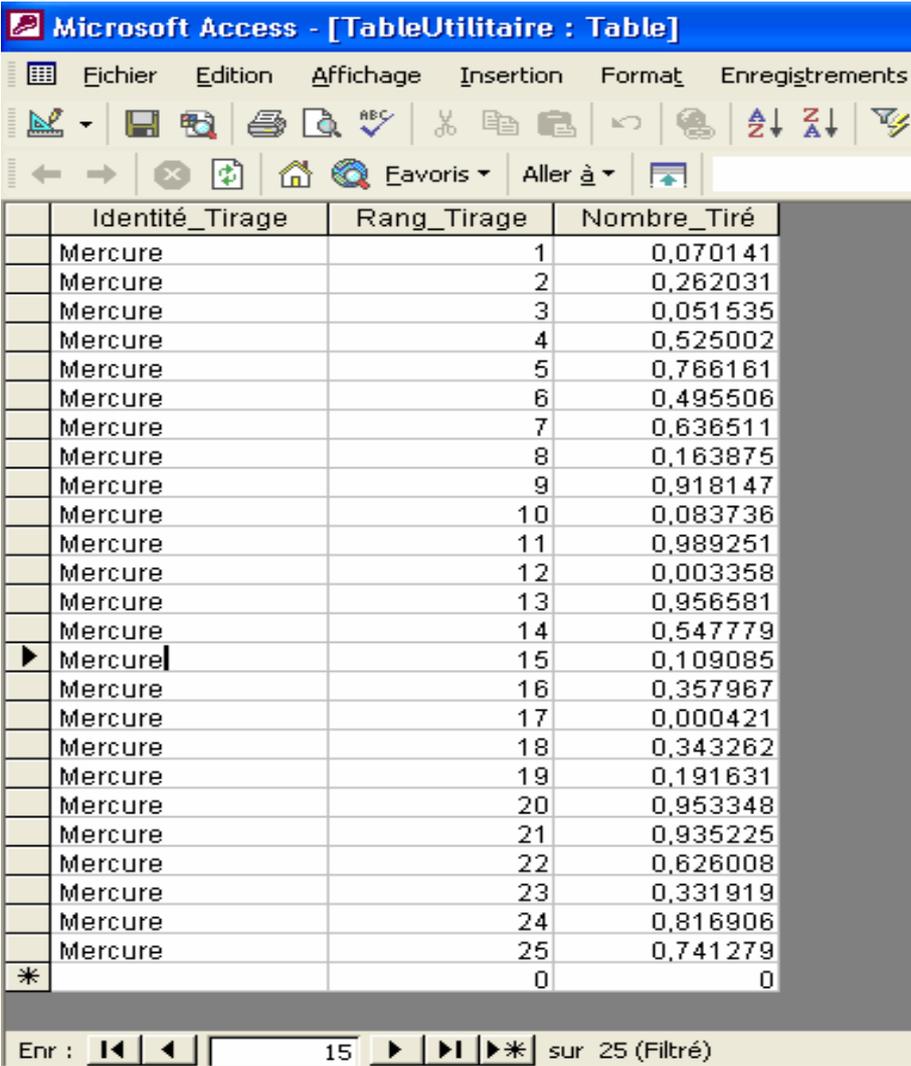
Fondements des méthodes de tirage utilisées

Ce chapitre 3 donne une brève présentation des fondements des méthodes de tirage utilisées pour certains des modules.

1. Principe de la génération des nombres aléatoires

Avec le logiciel Visual Basic, un nombre aléatoire se génère au moyen de la fonction Rnd à la suite de l'instruction Randomize qui initialise le générateur de nombres aléatoires à partir d'une valeur initiale tirée de l'horloge du système. L'exécution de la fonction Rnd fournit un nombre aléatoire suivant la loi uniforme sur l'intervalle [0,1[. A partir d'un tel nombre, on peut obtenir tout autre nombre aléatoire répondant à des conditions données. Quand on génère un ensemble de nombres aléatoires d'un même type, c'est-à-dire répondant aux mêmes conditions, on génère un échantillon aléatoire de nombres de ce type.

Table Access n° 11 : Un échantillon aléatoire de 25 nombres de l'intervalle [0,1[



The screenshot shows the Microsoft Access interface with a table named 'TableUtilitaire : Table'. The table has three columns: 'Identité_Tirage', 'Rang_Tirage', and 'Nombre_Tiré'. The data consists of 25 rows, each with the value 'Mercure' in the first column, a sequential number from 1 to 25 in the second column, and a random decimal value in the third column. The status bar at the bottom indicates 'Enr : 15 sur 25 (Filtré)'.

Identité_Tirage	Rang_Tirage	Nombre_Tiré
Mercure	1	0,070141
Mercure	2	0,262031
Mercure	3	0,051535
Mercure	4	0,525002
Mercure	5	0,766161
Mercure	6	0,495506
Mercure	7	0,636511
Mercure	8	0,163875
Mercure	9	0,918147
Mercure	10	0,083736
Mercure	11	0,989251
Mercure	12	0,003358
Mercure	13	0,956581
Mercure	14	0,547779
Mercure	15	0,109085
Mercure	16	0,357967
Mercure	17	0,000421
Mercure	18	0,343262
Mercure	19	0,191631
Mercure	20	0,953348
Mercure	21	0,935225
Mercure	22	0,626008
Mercure	23	0,331919
Mercure	24	0,816906
Mercure	25	0,741279
*	0	0

Le module "Générer un échantillon de nombres de l'intervalle [0,1[" du menu "Utilitaires" permet de générer des échantillons aléatoires de nombres de l'intervalle [0,1[. La table Access n° 11 ci-dessus fournit le résultat de la génération d'un échantillon aléatoire de 25 nombres de ce type.

On déduit de cet échantillon les échantillons aléatoires associés de nombres entiers respectivement de 1 chiffre, de 2 chiffres et de 3 chiffres. L'échantillon aléatoire de nombres de 1 chiffre est obtenu en multipliant les nombres du champ Nombre_Tiré de la table Access n°11 par 10 et en considérant la partie entière des résultats des multiplications. Les échantillons des nombres de 2 chiffres et de 3 chiffres sont obtenus en considérant la partie entière de la multiplication des nombres du champ Nombre_Tiré respectivement par 10^2 et 10^3 . Ces échantillons sont présentés dans le tableau n° 4 ci-dessous. On observe que 5 des 25 nombres de 1 chiffre sont le nombre 0 et qu'à ces 5 nombres, il ne correspond bien évidemment pas de nombres de 2 chiffres, ni de nombres de 3 chiffres. D'une manière générale, aux 25 nombres générés sur [0,1[, correspondent seulement 20 nombres entiers aléatoires associés de K chiffres pour $K > 1$.

Tableau n° 4 : Nombres de K chiffres associés à des nombres générés sur [0,1[

Nombre généralé sur [0,1[Nombre entier aléatoire associé			Nombre aléatoire associé de type texte de 5 chiffres	Nombre aléatoire entier associé compris entre 65 et 480
	de 1 chiffre	de 2 chiffres	de 3 chiffres		
0,070141	0			07014	70
0,262031	2	26	262	26203	262
0,051535	0			05153	
0,525002	5	52	525	52500	
0,766161	7	76	766	76616	
0,495506	4	49	495	49550	
0,636511	6	63	636	63651	
0,163875	1	16	163	16387	163
0,918147	9	91	918	91814	
0,083736	0			08373	83
0,989251	9	98	989	98925	
0,003358	0			00335	
0,956581	9	95	956	95658	
0,547779	5	54	547	54777	
0,109085	1	10	109	10908	109
0,357967	3	35	357	35796	357
0,000421	0			00042	
0,343262	3	34	343	34326	343
0,191631	1	19	191	19163	191
0,953348	9	95	953	95334	
0,935225	9	93	935	93522	
0,626008	6	62	626	62600	
0,331919	3	33	331	33191	331
0,816906	8	81	816	81690	
0,741279	7	74	741	74127	

2. Génération PEAR d'un échantillon de nombres de K chiffres de type texte

Le module "Génération PEAR de nombres de K chiffres de type texte" permet d'obtenir au hasard des nombres de K chiffres comme ceux d'une table de nombres au hasard, c'est-à-dire des nombres commençant éventuellement par un chiffre 0 ou plus. C'est ce type de nombre qui est appelé ici un nombre de type texte (ou alphanumérique). L'utilité de ce module est de générer, en cas de besoin, une table de nombres au hasard de K chiffres dont le fichier peut être reproduit facilement pour un usage sur le terrain dans une enquête.

La colonne 5 du tableau n° 4 ci-dessus présente l'échantillon PEAR de nombres de 5 chiffres de type de texte déduit de l'échantillon des 25 nombres aléatoires générés sur $[0,1[$.

3. Génération PEAR d'un échantillon de nombres entiers compris entre M et Z inclus

Soient M et Z deux nombres entiers avec $M < Z$, et soit L la longueur ou la taille du nombre entier Z, c'est-à-dire le nombre de chiffres dont Z est composé. Pour obtenir au hasard un nombre compris entre les entiers M et Z inclus, on doit générer un nombre compris entre 0 et 1 selon la loi uniforme sur $[0,1[$, multiplier ce nombre par 10^L et considérer la partie entière du résultat de la multiplication avec la condition que le nombre obtenu soit supérieur ou égal à M et inférieur ou égal à Z. Si cette double condition n'est pas réalisée, on génère un autre nombre sur $[0,1[$ et on poursuit cette opération autant de fois qu'il est nécessaire pour obtenir un nombre du type recherché.

Un échantillon PEAR de taille m de nombres de K chiffres compris entre M et Z s'obtient en générant m nombres remplissant cette condition et avec une possibilité de remise. A titre d'exemple, les 25 nombres aléatoires générés sur $[0,1[$ et qui figurent dans le tableau n° 4 ci-dessus donnent lieu à 9 nombres aléatoires entiers compris entre 65 et 480 comme le montre la colonne 6 de ce tableau.

4. Principe des tirages aléatoires

Soit U un univers composé de M unités et duquel on désire sélectionner au hasard un échantillon de m unités ou un échantillon de taille m. Les M unités statistiques de l'univers sont numérotées de 1 à M et ces numéros constituent les identités ou étiquettes des unités. Tirer un échantillon aléatoire de taille m de l'univers U consistera simplement à générer, à l'aide de la fonction Rnd de Visual Basic, m nombres compris entre 1 et M qui identifieront les unités sélectionnées. La génération des m nombres tiendra compte bien évidemment des spécificités du mode particulier de tirage considéré.

Tous les tirages utilisés dans TIRAGE 2.1 sont fondés sur ce principe de base même si certains tirages comme les tirages systématiques sont des cas légèrement plus complexes de mise en œuvre de ce principe.

5. Tirage PIAR d'un échantillon d'unités statistiques au sein d'une population

Soit U un univers composé de M unités statistiques d'étiquettes 1, 2, ..., r, ..., M et de tailles respectives $X_1, X_2, \dots, X_r, \dots, X_M$. La taille est souvent l'effectif des exploitations agricoles, l'effectif de la population ou l'effectif des ménages lorsque les unités statistiques sont, par exemple, les villages d'une région.

Un cas classique de tirage PIAR (tirage avec probabilités inégales et avec remise) d'un échantillon d'unités au sein d'un tel univers consiste à prendre les probabilités inégales proportionnelles aux tailles X_r des unités statistiques.

Soit C_r la somme des tailles des r premières unités statistiques $1, 2, \dots, r$ de l'univers U . C_r est également appelé le cumul des tailles des r premières unités. Il est défini par la relation

$$C_r = \sum_{j=1}^r X_j$$

Nous désignerons par C le cumul total des tailles ou le cumul des tailles pour l'ensemble des unités de l'univers U , soit

$$C = C_M = \sum_{j=1}^M X_j$$

Pour mettre en œuvre le tirage PIAR, on établit au préalable une partition de l'ensemble des nombres entiers $\{1, 2, \dots, C\}$ en M éléments et on associe l'unité r de l'univers à l'élément r de la partition défini par le sous-ensemble des nombres entiers $\{1+C_{r-1}, \dots, C_r\}$.

Le tableau ci-après présente les caractéristiques de la partition. On note que par construction des grandeurs C_h , l'effectif des nombres entiers de l'élément r de la partition est égal à la taille X_r .

<u>Unité de l'univers</u>	<u>Élément de la partition</u>	<u>Effectif des nombres de la partition</u>
1	$\{1, 2, \dots, C_1\}$	X_1
2	$\{1 + C_1, \dots, C_2\}$	X_2
.....
r	$\{1 + C_{r-1}, \dots, C_r\}$	X_r
.....
M	$\{1 + C_{M-1}, \dots, C_M\}$	X_M

Sélectionner une unité de l'univers U consiste alors à tirer un des M éléments de la partition avec des probabilités proportionnelles aux tailles X_r de ces éléments. Cela revient à tirer au hasard, avec probabilités égales, un nombre compris entre 1 et C , puis à identifier ce nombre avec l'élément de la partition et l'unité de l'univers qui lui correspondent.

Pour obtenir l'échantillon PIAR de taille m , on répète le procédé précédent m fois avec une possibilité de remise pour les tirages.

Les tirages PIAR auxquels donne lieu le logiciel TIRAGE 2.1 sont basés sur cette méthode. Ils requièrent que la table qui sert de source de données pour les tirages comporte un champ Cumul avec les valeurs C_r des cumuls des tailles des unités. La création de ce champ ainsi que le calcul des valeurs de ce champ sont exécutés directement par le module "Tirage PIAR d'un échantillon d'unités au sein d'une population".

Exemple d'application

Considérons la table ABCD de la base de données Data.mdb. ABCD est une base de sondage constituée de 20 villages imaginaires dont le nombre de ménages est connu pour chacun. Le

nombre de ménages joue le rôle de taille du village. Tirons un échantillon de 4 villages selon le mode PIAR de cette base de sondage. Le tableau n° 5 ci-dessous contient les données relatives au tirage réalisé.

Tableau n° 5 : Données d'un tirage PIAR de 4 unités de l'univers ABCD

Identité du village	Nom du village	Effectif des ménages	Cumul des tailles	Nombre tiré	Unité tirée
8121	Assabe	136	136		
8122	Atal	240	376		
8123	Gorgui	115	491		
8124	Kaka	127	618	515	x
8125	Kaporo	181	799		
8126	Kipé	184	983		
8127	Lalo	235	1 218		
8128	Loum	277	1 495	1 241	x
8129	Madina	211	1 706	1 541	x
8130	Manga	112	1 818		
8131	Mano	210	2 028		
8132	Marsa	143	2 171		
8133	Matoto	223	2 394		
8134	Saga	252	2 646		
8135	Sindaya	266	2 912		
8136	Taouya	102	3 014		
8137	Tilen	126	3 140	3 090	x
8138	Valo	203	3 343		
8139	Vanta	148	3 491		
8140	Vival	149	3 640		

Le cumul total C prend la valeur 3 640 et les 4 nombres aléatoires tirés avec remise entre 1 et 3 640 inclus sont : 515, 1241, 1541 et 3090.

Avec chacun de ces nombres tirés, comment identifier l'unité correspondante de l'univers ? Si Z est l'un des nombres tirés, on calcule successivement la différence $Z - C_r$ en parcourant les valeurs de C_r selon l'ordre croissant jusqu'à ce que la différence devienne négative. Le premier cumul qui rend négative cette différence est le cumul de l'unité correspondant au nombre Z tiré. On observe selon le tableau n° 5 que les 4 nombres aléatoires générés 515, 1241, 1541 et 3090 correspondent respectivement aux cumuls 618, 1495, 1706 et 3140. Les nombres tirés sont inscrits dans la colonne 5 de ce tableau en face du cumul C_r de l'unité r correspondante. Les unités tirées sont marquées d'une croix dans la colonne 6. Les villages de Kaka, Loum, Madina et Tilen constituent l'échantillon tiré.

6. Tirage SPE (tirage systématique avec probabilités égales) d'un échantillon d'unités

Le tirage SPE ou tirage systématique avec probabilités égales est une variante du tirage aléatoire simple sans remise, c'est-à-dire du tirage PESR. Si la constitution des échantillons conduit à des résultats équivalents pour les deux méthodes, ces dernières sont cependant différentes parce que correspondant à deux modèles probabilistes différents.

Soit U un univers de M unités statistiques d'étiquettes 1, 2, ..., r, ..., M. L'on désire tirer de cet univers, un échantillon de taille m avec probabilités égales et sans remise selon le tirage SPE. Supposons M divisible par m avec

$$M = mk$$

k étant donc un nombre entier. Pour mettre en œuvre le tirage SPE, on répartit au préalable les unités de l'univers U en m groupes de taille k chacun. La méthode consiste alors à tirer avec probabilités égales, une unité de chacun des m groupes d'unités. On commence par tirer une unité du groupe n°1 constituée par l'ensemble {1, 2, ..., k}. Si r est l'étiquette de l'unité tirée, avec $r \leq k$, les m-1 unités à retenir dans les autres groupes sont ceux dont les étiquettes sont les termes suivants de la progression arithmétique finie de raison k et de premier terme r.

$$r, r+k, r+2k, \dots, r+(m-1)k .$$

On parle de tirage systématique de raison k. Dans le cas où M est divisible par m, toutes les unités ont la même probabilité d'être tirée ou d'appartenir à l'échantillon.

Dans le cas où M n'est pas divisible par m et où k n'est donc pas un nombre entier, on ne peut pas constituer des groupes d'unités de même taille comme précédemment. Pour réaliser le tirage systématique avec probabilités égales, on adapte la méthode précédente en générant un nombre u selon la loi uniforme sur $[0, 1[$ et on prend pour première unité tirée, celle qui a pour étiquette, le nombre $[ku]+1$, où $[ku]$ est la partie entière de ku. Et l'échantillon tout entier est donné comme suit :

<u>Rang du tirage</u>	<u>Nombre généralisé</u>
1	$[ku] + 1$
2	$[ku + k] + 1$
3	$[ku + 2k] + 1$
.....	
r	$[ku + (r-1)k] + 1$
.....	
m	$[ku + (m-1)k] + 1$

Tout revient à considérer une partition de l'ensemble $\{1, 2, 3, \dots, M\}$ en m éléments qui sont les intervalles $]0, k]$, $]k, 2k]$, ..., $](r-1)k, rk]$, ..., $](m-1)k, M]$ dont on considère les seuls nombres entiers. L'unité de rang r de l'échantillon, soit le nombre $[ku+(r-1)k] + 1$, est issue de l'élément $](r-1)k, rk]$ de la partition.

Le module "Tirage systématique avec probabilités égales d'un échantillon d'unités" est fondé sur cette dernière procédure qui s'applique également au cas où M est divisible par m.

Exemple d'application 1

Considérons la population $U = \{1, 2, \dots, 52\}$ et sélectionnons 6 unités de cette population selon le mode de tirage SPE. Nous avons $m = 6$, $k = 8,66$ et $u = 0,738$.

Tableau n° 6 : Données d'un tirage SPE de 6 unités de la population $\{1, 2, \dots, 52\}$

Rang du tirage r	$ku+(r-1)k$	$[ku+(r-1)k]+1$
1	6,39108	7
2	15,05108	16
3	23,71108	24
4	32,37108	33
5	41,03108	42
6	49,69108	50

Les 6 éléments de la partition de l'ensemble des nombres $\{1, 2, \dots, 52\}$ sont les sous-ensembles $\{1, 2, \dots, 8\}$, $\{9, 10, \dots, 17\}$, $\{18, 19, \dots, 26\}$, $\{27, 28, \dots, 34\}$, $\{35, 36, \dots, 43\}$ et $\{44, 45, \dots, 52\}$. Les autres données du tirage sont fournies par le tableau n° 6 ci-dessus.

Exemple d'application 2

Sélectionnons selon le mode de tirage SPE, un échantillon de 3 villages de la région des 20 villages représentée par la table ABCD de la base de données Data.mdb. Les données de départ du tirage sont :

$$\begin{aligned} M &= 20 \\ m &= 3 \\ u &= 0,3914912 \end{aligned}$$

Le tableau n° 7 ci-dessous contient les données du tirage SPE. Les villages de numéros 3, 10 et 16 ont été tirés, c'est-à-dire les villages Gorgui, Manga et Taouya.

Tableau n° 7 : Données d'un tirage SPE de 3 villages de la région ABCD

Rang du tirage r	$ku+(r-1)k$	$[ku+(r-1)k]+1$
1	2,60994	3
2	9,27661	10
3	15,94327	16

7. Tirage systématique avec probabilités proportionnelles à la taille des unités : une version simplifiée de la procédure de Hartley et Rao

Les sondages avec probabilités inégales et sans remise mis en œuvre sont généralement des sondages utilisant la méthode du tirage systématique. Il existe de nombreuses procédures de tirage systématique avec probabilités inégales.

Soit U une population composée des M unités statistiques d'étiquettes $1, 2, \dots, r, \dots, M$. Généralement, on définit la probabilité inégale P_r de tirage de l'unité d'étiquette r à partir d'une variable auxiliaire X_r connue pour toutes les unités de l'univers U et à peu près proportionnelle à la variable d'étude Y_r du sondage mis en œuvre. Si P_r est donc exactement proportionnel à X_r et que X_r est à peu près proportionnel à Y_r , on aura P_r à peu près proportionnel à Y_r .

Bien souvent, la variable auxiliaire X_r est une variable de taille des unités. C'est pourquoi on parle dans ce cas, de tirage systématique avec probabilités proportionnelles à la taille des unités. Nous présentons ici la méthode de tirage systématique de Hartley et Rao⁶, une procédure de tirage systématique parmi les plus simples.

⁶ H. O. Hartley et J. K. N. Rao, Sampling with unequal probabilities without replacement, Annals of Mathematical Statistics, 1962, vol. 33

Soit U une population composée des M unités statistiques d'étiquettes 1, 2, ..., r, ..., M et de tailles respectives $X_1, X_2, \dots, X_r, \dots, X_M$. La taille des unités est une variable auxiliaire que l'on suppose connue pour toutes les unités de la population. On désire tirer un échantillon de m unités de la population U selon la procédure de tirage de Hartley et Rao.

Désignons par C la somme cumulée totale des tailles X_r , soit

$$C = \sum_{j=1}^M X_j$$

Le cumul total C doit être divisible par m, la taille de l'échantillon. Si tel n'est pas le cas, on multiplie toutes les tailles X_r par une constante de manière à rendre C divisible par m. On prend souvent cette constante égale à m. Cette transformation des tailles X_r et du cumul total C en les multipliant par m ne change rien à la procédure de tirage.

Le cumul total C étant divisible par m, il vient

$$C = m K$$

où K est un nombre entier. L'ensemble des nombres entiers $\{1, 2, \dots, C\}$ est alors subdivisé en m sous-ensembles comportant chacun K nombres.

Tirer m unités de la population U selon le tirage systématique avec probabilités proportionnelles aux tailles X_r , revient à tirer avec probabilités égales et sans remise, un nombre entier dans chacun des m sous-ensembles de K chiffres constitués et à identifier l'unité correspondant à ce nombre. Pour ce faire, on génère au départ un nombre entier compris entre 1 et K inclus. Si r est le nombre généré, les m-1 autres nombres tirés des m-1 autres sous-ensembles de nombres constitués sont les m-1 termes suivants de la suite arithmétique finie de raison K et de premier terme r. L'échantillon des m nombres tirés est alors constitué par l'ensemble $\{r, r+K, \dots, r+(m-1)K\}$. Chacun de ces m nombres tirés peut s'identifier à une unité de la population et à une seule.

En réalité, chacun des m sous-groupes donne lieu au tirage d'une seule unité si $K > X_r$ pour tout $r = 1, 2, \dots, M$. Si pour une unité r, $K < X_r$, alors cette unité risque d'être tirée plus d'une fois. On décide alors, soit de sélectionner d'office cette unité avec la probabilité 1 et de tirer m-1 autres unités parmi les autres M-1 unités restantes, soit de segmenter cette unité en 2 nouvelles unités ou plus avant les tirages.

Désignons par C_r le cumul des tailles pour les r premières unités de la population, soit

$$C_r = \sum_{j=1}^r X_j \quad \forall r = 1, 2, \dots, M$$

Notons que

$$C = C_M = \sum_{j=1}^M X_j$$

Comment identifier l'unité de la population qui correspond au nombre r tiré entre 1 et K ? On calcule successivement la différence $r - C_h$ en parcourant selon l'ordre croissant, les cumuls C_h jusqu'à ce que la différence devienne négative. Le premier cumul qui rend négative cette différence est le cumul de l'unité correspondant au nombre r tiré. Le nombre $r+(j-1)K$ tiré du sous-ensemble $\{1+(j-1)K, \dots, jK\}$ correspond lui à l'unité dont le cumul C_h rend négative le premier, la différence $r+(j-1)K - C_h$ lorsque ces cumuls sont parcourus dans l'ordre croissant.

Exemple d'application

Sélectionnons selon le mode de tirage SPPT, un échantillon de 3 villages de la région des 20 villages représentée par la table ABCD de la base de données Data.mdb. La taille des villages est le champ Nombre_Ménages.

Les tailles X_r sont multipliées par la taille de l'échantillon, soit par 3. Et nous obtenons les données de départ suivantes :

$$m = 3$$

$$C = 10920$$

$$K = 3640$$

Nous observons que la condition $K > X_r$ est vérifiée pour tout $r = 1, 2, \dots, M$. La procédure de tirage s'applique donc. Les résultats du tirage sont résumés ci-après tandis que le tableau n° 8 fournit l'ensemble des données du tirage.

<u>Nombre tiré</u>	<u>Sous-ensemble</u>	<u>Cumul associé</u>	<u>Village tiré</u>
3293	{1, 2, , 3640}	3654	Lalo
6933	{3641, , 7280}	7182	Matoto
10573	{7281, , 10920}	10920	Vival

Tableau n° 8 : Données d'un tirage SPPT de 3 villages de la région ABCD

Numéro d'ordre	Identité du village	Nom du village	Nombre de ménages	Cumul des tailles	Cumul modifié des tailles	Nombre tiré	Unité tirée
1	8121	Assabe	136	136	408		
2	8122	Atal	240	376	1 128		
3	8123	Gorgui	115	491	1 473		
4	8124	Kaka	127	618	1 854		
5	8125	Kaporo	181	799	2 397		
6	8126	Kipé	184	983	2 949		
7	8127	Lalo	235	1 218	3 654	3 293	x
8	8128	Loum	277	1 495	4 485		
9	8129	Madina	211	1 706	5 118		
10	8130	Manga	112	1 818	5 454		
11	8131	Mano	210	2 028	6 084		
12	8132	Marsa	143	2 171	6 513		
13	8133	Matoto	223	2 394	7 182	6 933	x
14	8134	Saga	252	2 646	7 938		
15	8135	Sindaya	266	2 912	8 736		
16	8136	Taouya	102	3 014	9 042		
17	8137	Tilen	126	3 140	9 420		
18	8138	Valo	203	3 343	10 029		
19	8139	Vanta	148	3 491	10 473		
20	8140	Vival	149	3 640	10 920	10 573	x

8. Tirage systématique avec probabilités proportionnelles à la taille des unités : la version originale de la procédure de Hartley et Rao

Pour des besoins d'estimation, la méthode de Hartley et Rao requiert que les unités de la base de sondage soient triées de façon aléatoire, c'est-à-dire réordonnées au hasard avant la mise en œuvre du tirage.

A titre d'exemple, appliquons la méthode de tirage complète de Hartley et Rao pour sélectionner 3 villages parmi les 20 villages de la région ABCD. Nous commençons par réordonner les villages de ABCD à l'aide du module "Réordonner au hasard les unités d'une population".

Les colonnes 1 et 2 du tableau n° 9 ci-dessous fournissent respectivement l'ancien numéro d'ordre (nombre tiré) et le nouveau numéro d'ordre des villages de la région ABCD. L'identité du village, le nom du village et la taille du village sont désormais dans un autre ordre différent de celui du départ. Il s'agit de l'ordre défini par le nouveau numéro d'ordre du village à la colonne 2 du tableau n° 9. Il en résulte que le champ Cumul voit ses valeurs changées comme cela apparaît quand on compare les tableaux n° 8 et n° 9.

Comme pour le tirage précédent, les tailles des unités sont multipliées par m. Nous avons les données de départ suivantes :

$$m = 3, \quad C = 10\,920, \quad K = 3\,640 \text{ et } K > X_r, \quad \forall r = 1, 2, \dots, 20$$

Les résultats du tirage sont résumés ci-après et le tableau n° 9 fournit l'ensemble des données du tirage.

<u>Nombre tiré</u>	<u>Sous-ensemble</u>	<u>Cumul associé</u>	<u>Village tiré</u>
2233	{1, 2, , 3640}	2487	Kaporo
5873	{3641, , 7280}	5948	Sindaya
9513	{7281, , 10920}	9645	Kipé

Tableau n° 9 : Données d'un tirage de 3 villages de ABCD selon la méthode complète de Hartley et Rao

Nombre tiré	Nouveau numéro d'ordre	Identité du village	Nom du village	Taille du village	Cumul des tailles	Cumul modifié des tailles	Nombre tiré	Unité tirée
10	1	8130	Manga	112	112	336		
13	2	8133	Matoto	223	335	1 005		
9	3	8129	Madina	211	546	1 638		
16	4	8136	Taouya	102	648	1 944		
5	5	8125	Kaporo	181	829	2 487	2 233	x
12	6	8132	Marsa	143	972	2 916		
3	7	8123	Gorgui	115	1 087	3 261		
17	8	8137	Tilen	126	1 213	3 639		
1	9	8121	Assabe	136	1 349	4 047		
4	10	8124	Kaka	127	1 476	4 428		
2	11	8122	Atal	240	1 716	5 148		
15	12	8135	Sindaya	266	1 982	5 946	5 873	x
7	13	8127	Lalo	235	2 217	6 651		
20	14	8140	Vival	149	2 366	7 098		
18	15	8138	Valo	203	2 569	7 707		
14	16	8134	Saga	252	2 821	8 463		
11	17	8131	Mano	210	3 031	9 093		
6	18	8126	Kipé	184	3 215	9 645	9 513	x
8	19	8128	Loum	277	3 492	10 476		
19	20	8139	Vanta	148	3 640	10 920		

Remarque

Le tirage précédent effectué selon la méthode complète de Hartley et Rao a été réalisé au moyen d'une feuille Excel. Si l'on veut utiliser le module "Tirage SPPT d'un échantillon d'unités", il sera appliqué non pas à la table ABCD mais plutôt à la table TableOrdre qui contient les transformations apportées à l'ordre des villages par le module "Réordonner au hasard les unités d'une population".

Ensuite, lors du tirage, le champ "Identité de l'unité statistique" qui devra être choisi est le nouveau numéro d'ordre de l'unité, autrement le tirage sera réalisé selon l'ordre ancien des unités.

9. Procédure de tirage SPPT basée sur les probabilités d'inclusion d'ordre 1

La procédure du tirage SPPT peut être formulée autrement au moyen des probabilités d'inclusion des unités de la population sondée. Soit U une population composée des M unités statistiques d'étiquettes $1, 2, \dots, r, \dots, M$ et de tailles respectives $X_1, X_2, \dots, X_r, \dots, X_M$. On désire sélectionner m unités de la population U selon le mode de tirage SPPT.

On désigne par π_r la probabilité d'inclusion d'ordre 1 de l'unité r , c'est-à-dire la probabilité pour l'unité r de faire partie de l'échantillon des m unités.

Désignons par X la somme cumulée totale des tailles X_r , soit

$$X = \sum_{j=1}^M X_j$$

La taille relative p_r de l'unité r est définie par la relation

$$p_r = X_r / X \quad \forall r = 1, 2, \dots, M$$

Pour un tirage SPPT, la probabilité d'inclusion π_r est proportionnelle à la taille relative p_r , ce qui s'exprime par la double relation

$$\pi_r = k p_r = k X_r / X \quad \forall r = 1, 2, \dots, M$$

où k est une constante.

On montre que pour tout plan de sondage de taille fixe m (taille de l'échantillon), on a

$$\sum_{r=1}^M \pi_r = m$$

On en déduit

$$m = \sum_{r=1}^M \pi_r = k \sum_{r=1}^M \frac{X_r}{X} = k$$

D'où

$$\pi_r = m p_r = m X_r / X \quad \forall r = 1, 2, \dots, M$$

La procédure de tirage fondée sur les probabilités d'inclusion exige que soit vérifiée la condition :

$$0 < \pi_r < 1 \quad \forall r = 1, 2, \dots, M$$

Si pour une unité r , $\pi_r > 1$, ou bien cette unité est désignée d'office pour faire partie de l'échantillon avec la probabilité 1 et on aura alors à tirer $m-1$ autres unités parmi les $M - 1$ unités restantes, ou bien l'unité est segmentée en 2 nouvelles unités ou plus avant les tirages.

La procédure de tirage est fondée sur les cumuls des probabilités d'inclusion à l'instar de la procédure de tirage fondée sur les cumuls des tailles des unités. Désignons par W_r le cumul des probabilités d'inclusion pour les r premières unités de la population U , soit

$$W_r = \sum_{j=1}^r \pi_j \quad \forall r = 1, 2, \dots, M$$

Notons que le cumul total W ou W_M des probabilités d'inclusion vérifie la relation

$$W = W_M = \sum_{j=1}^M \pi_j = m$$

Pour réaliser le tirage, on commence par générer un nombre aléatoire selon la loi uniforme sur l'intervalle $[0,1]$. Si u est le nombre obtenu, les m unités de l'échantillon, soit r_1, r_2, \dots, r_m , sont associés aux m nombres $u, u+1, \dots, u+(m-1)$ selon les conditions ci-après.

La première unité r_1 vérifiera la condition

$$W_{r_1-1} \leq u < W_{r_1}$$

La deuxième unité r_2 , l'unité r_h et l'unité r_m vérifieront respectivement les conditions

$$W_{r_2-1} \leq u+1 < W_{r_2}$$

$$W_{r_h-1} \leq u+(h-1) < W_{r_h}$$

et $W_{r_m-1} \leq u+(m-1) < W_{r_m}$

Tableau n° 10 : Données d'un tirage SPPT de 3 villages de ABCD fondé sur les cumuls des probabilités d'inclusion

Numéro d'ordre r	Identité du village	Nom du village	Taille de l'unité X_r	Taille relative de l'unité p_r	Probabilité d'inclusion d'ordre 1 π_r	Cumul des probabilités d'inclusion W_r	Nombre généré	Unité tirée
1	8121	Assabe	136	0,03736	0,11209	0,11209		
2	8122	Atal	240	0,06593	0,19780	0,30989		
3	8123	Gorgui	115	0,03159	0,09478	0,40467		
4	8124	Kaka	127	0,03489	0,10467	0,50934		
5	8125	Kaporo	181	0,04973	0,14918	0,65852	0,61220	x
6	8126	Kipé	184	0,05055	0,15165	0,81016		
7	8127	Lalo	235	0,06456	0,19368	1,00385		
8	8128	Loum	277	0,07610	0,22830	1,23214		
9	8129	Madina	211	0,05797	0,17390	1,40604		
10	8130	Manga	112	0,03077	0,09231	1,49835		
11	8131	Mano	210	0,05769	0,17308	1,67143	1,61220	x
12	8132	Marsa	143	0,03929	0,11786	1,78929		
13	8133	Matoto	223	0,06126	0,18379	1,97308		
14	8134	Saga	252	0,06923	0,20769	2,18077		
15	8135	Sindaya	266	0,07308	0,21923	2,40000		
16	8136	Taouya	102	0,02802	0,08407	2,48407		
17	8137	Tilen	126	0,03462	0,10385	2,58791		
18	8138	Valo	203	0,05577	0,16731	2,75522	2,61220	x
19	8139	Vanta	148	0,04066	0,12198	2,87720		
20	8140	Vival	149	0,04093	0,12280	3,00000		

Exemple d'application

Sélectionnons selon le mode de tirage SPPT fondé sur les cumuls des probabilités d'inclusion d'ordre 1, un échantillon de 3 villages de la région des 20 villages représentée par la table ABCD de la base de données Data.mdb.

Nous obtenons les données de départ suivantes :

$$\begin{aligned} m &= 3 \\ u &= 0,61220 \\ 0 < \pi_r < 1 & \quad \forall r = 1, 2, \dots, M \end{aligned}$$

Les autres données du tirage sont présentées dans la tableau n° 10 ci-dessus. Nous notons que ce dernier tableau est simplement une adaptation du tableau n° 8 relatif au tirage fondé sur les cumuls des tailles des unités.

Comment identifier l'unité r_1 de la population U qui correspond au nombre u ? On calcule successivement la différence $u - W_h$ en parcourant selon l'ordre croissant, les cumuls W_h jusqu'à ce que la différence devienne négative. Le premier cumul qui rend négative cette différence est le cumul de l'unité correspondant au nombre r_1 tiré, soit l'unité de numéro d'ordre 5, c'est-à-dire le village Kaporo. De même, les nombres r_2 et r_3 tirés correspondent respectivement aux cumuls qui rendent négatives les différences $u+1 - W_h$ et $u+2 - W_h$ les premiers lorsque ces cumuls sont parcourus dans l'ordre croissant. Il leur correspond respectivement les unités de numéros d'ordre 11 et 18, soit les villages Mano et Valo.

Remarque

Les deux formulations reposant sur le cumul des tailles ou le cumul des probabilités d'inclusion d'ordre 1 sont équivalentes. On observe notamment, étant donné que

$$K = X/m$$

que la condition $\pi_r < 1 \quad \forall r = 1, 2, \dots, M$

entraîne la condition $\pi_r = m X_r / X = X_r / K < 1 \quad \forall r = 1, 2, \dots, M$

D'où la condition $X_r < K \quad \forall r = 1, 2, \dots, M$

qui est la condition à laquelle les tailles doivent obéir dans la procédure fondée sur le cumul des tailles des unités.