

# Exercice 1: Comprendre la dynamique des populations

Visseho Adjiwanou, PhD.

07 June 2021

Cet exercice est tiré du livre de Imai Kosuke.

Comprendre la dynamique des populations est important pour de nombreux domaines des sciences sociales. Nous allons calculer certaines quantités démographiques à base des naissances et des décès pour deux pays et la population mondiale pour deux périodes: 1950 à 1955 et 2005 à 2010. Nous analyserons les fichiers de données CSV suivants: `Kenya.csv`, `Sweden.csv` et `World.csv`. Chaque fichier contient des données démographiques pour le Kenya, la Suède (Sweden) et le monde (World), respectivement. Le tableau ci-dessous présente les noms et les descriptions des variables dans chaque ensemble de données.

Nom	Description
<code>country</code>	Nom du pays abrégé
<code>period</code>	Période au cours de laquelle les données sont collectées
<code>age</code>	Groupe d'âge
<code>births</code>	Nombre de naissance en milliers (i.e., nombre d'enfants nés des femmes du groupe d'âge)
<code>deaths</code>	Nombre de décès en millier
<code>py.men</code>	Personne-années pour les hommes en milliers
<code>py.women</code>	Personne-années pour les femmes en milliers

Source: United Nations, Department of Economic and Social Affairs, Population Division (2013). *World Population Prospects: The 2012 Revision, DVD Edition*.

Les données sont collectées pour une période de 5 ans, où *personne-années* est une mesure de la contribution temporelle de chaque personne au cours de la période. Par exemple, une personne qui vit pendant toute la période de cinq ans contribue pour cinq personnes-années alors que celle qui ne vit que pendant la première moitié de la période contribue à 2,5 personnes-années. Avant de commencer cet exercice, il serait judicieux d'inspecter directement chaque ensemble de données. Dans R, cela peut être fait avec la fonction `View`, qui prend comme argument le nom d'un fichier (`data.frame`) à examiner. Alternativement, dans RStudio, double-cliquer sur le nom du fichier que vous venez d'ouvrir dans l'onglet **Environment** vous permettra de visualiser les données dans une vue de type tableur. Vous pouvez directement visualiser les données dans RMarkdown avec la commande `head`.

**Remarque 1:** Prenez toujours soins d'enregistrer les **données à utiliser dans le même fichier** que votre RMarkdown pour vous simplifier la vie.

**Remarque 2:** Tout au long de l'exercice, pensez à écrire de petite fonctions pour simplifier votre tâche.

## Mise en route

Ce code vous permet de charger directement les données. Vous devez télécharger les données de Moodle et les placer avec ce fichier Rmarkdown dans le même dossier (voir remarque en haut).

```
## Telecharger les données

Sweden <- read.csv("../Données/Sweden.csv")
Kenya <- read.csv("../Données/Kenya.csv")
World <- read.csv("../Données/World.csv")

#head(Sweden)
```

## QUESTION 1

Nous commençons par calculer *le taux brut de natalité* (TBN) pour une période donnée. Le TBN est défini comme suit:

$$\text{TBN} = \frac{\text{Nombre total de naissances}}{\text{Nombre total de personnes-années vécues}}$$

Calculez le TBN pour chaque période, séparément pour le Kenya, la Suède et le Monde. Pour faire ceci, il vous faut créer peut-être plusieurs objets: 1. NTNKP1 signifie : nombre total de naissance du Kenya pour la période 1. 2. NTAVHKP1 signifie Nombre total des années vécues par les hommes au Kenya pour la période 1 et ainsi de suite. Je vous donne ces résultats dans la commande.

### Réponse 1

Votre interprétation

## QUESTION 2

Prenez un moment pour réfléchir à la manière d'organiser les résultats précédents.

Stockez les résultats du TBN sous forme de matrice de dimension (2, 3), 2 étant le nombre de période et 3 le nombre de pays/région. Décrivez brièvement les tendances observées dans les TBN calculés.

### Réponse 2

Votre interprétation

## QUESTION 3

Le TBN est facile à comprendre mais pose des problèmes de comparaisons. Pourquoi le taux brut de natalité n'est pas un bon indicateur de comparaison du niveau de fécondité entre pays?

### Réponse 3

## QUESTION 4

Une solution intermédiaire au TBN est le calcul du *taux général de fécondité* (TGF). Celui-ci se calcule par:

$$\text{TGF} = \frac{\text{Nombre total de naissances}}{\text{Nombre total de personnes-années vécues par les femmes de 15-49 ans}}$$

Calculer ce taux pour le Kenya, la Suède et le Monde. Présenter les résultats aussi sous forme de matrice. Pourquoi cela vous paraît déjà meilleur que le TBN? Interpréter les résultats. Quel est le problème du taux général de fécondité.

## Réponse 4

### Votre interprétation

## QUESTION 5

Pour remédier à ces problèmes, nous calculons l'*indice synthétique de fécondité* (ISF). Contrairement au TBN, l'ISF s'ajuste en fonction de la composition par âge de la population féminine en âge de procréer. Pour ce faire, nous devons d'abord calculer le *taux de fécondité par âge* (TFA), qui représente le taux de fécondité chez les femmes en âge de procréer  $[15, 50]$ . Le TFA pour la tranche d'âge  $[x, x + \delta)$ , où  $x$  est l'âge de départ et  $\delta$  la largeur de la tranche d'âge (mesurée en années), est défini comme suit:

$$\text{TFA}_{[x, x+\delta)} = \frac{\text{Nombre de naissances des femmes âgées de } [x, x + \delta)}{\text{Nombre de personnes-années vécues par les femmes âgées } [x, x + \delta)}$$

Notez que les crochets,  $[ , ]$ , incluent la limite alors que les parenthèses,  $( , )$ , l'excluent. Par exemple,  $[20, 25)$  représente la tranche d'âge supérieure ou égale à 20 ans et inférieure à 25 ans. Dans les données démographiques typiques, la tranche d'âge  $\delta$  est fixée à 5 ans.

Calculez le taux de fécondité par âge pour la Suède et le Kenya ainsi que pour le monde entier pour chacune des deux périodes. Stockez les TFA résultants séparément pour chaque région. Interpréter ces résultats. Comment devriez-vous présenter ces résultats pour une interprétation optimale?

## Réponse 5

### Votre interprétation

## QUESTION 6

En utilisant le TFA, nous pouvons définir l'ISF comme le nombre moyen d'enfants que les femmes mettent au monde si elles vivent tout au long de leur vie reproductive.

$$\text{ISF} = \text{TFA}_{[15, 20)} \times 5 + \text{TFA}_{[20, 25)} \times 5 + \cdots + \text{TFA}_{[45, 50)} \times 5$$

Nous multiplions chaque taux de fécondité par âge par 5 car la tranche d'âge est de 5 ans. Calculez l'ISF pour la Suède et le Kenya ainsi que le monde entier pour chacune des deux périodes. Comme dans la question précédente, continuer à supposer que la tranche d'âge en âge de procréer des femmes est de  $[15, 50)$ . Stockez les deux ISF résultants pour chaque pays et le monde en tant que matrice de dimension  $(2, 3)$ . En général, comment est-ce que l'indice synthétique de fécondité a-t-il évolué au cours de la période de 1950 à 2000 dans le monde? Commentez les résultats de la Suède et du Kenya.

## Réponse 6

### Votre interprétation

# ANNEXES

## 1. Glossaire

Fécondité: C'est l'étude des naissances

Fécondité = Fertility

Fécondité  $\neq$  Fertilité

Taux brut de natalité = Crude birth rate (CBR)

Taux global de fécondité générale

Indice synthétique de fécondité (ISF) = Total fertility rate (TFR)

Taux de fécondité par âge = Age specific fertility rate (ASFR)