

Macro-projection de population vs micro-simulation

1. — Introduction

Les méthodes de micro-simulation ont d'abord été développées comme une alternative aux méthodes de macro-projection de population. Les macro-projection de population se donnent comme objet la projection, typiquement pour les cinquante ans à venir, de la pyramide des âges de la population d'un pays. Ces méthodes utilisent la pyramide des âges initiale et recourent à des hypothèses sur les évolutions, respectivement, de la mortalité par âge et de la fécondité des femmes en âge d'avoir des enfants. Notons N_{gt}^f les effectifs des femmes d'âge g de l'année t . Le nombre de femmes, en $t+1$, d'âge $g+1$ se calcule à partir de l'équation :

$$N_{g+1,t+1}^f = ts_{gt}^f \times N_{gt}^f$$

où ts_{gt}^f est le taux de survie des femmes d'âge g de l'année t .

Le nombre total de naissances dépend du nombre de femmes en âge d'avoir des enfants et des taux de fécondité à chaque âge :

$$N_{0,t+1} = \sum_{g=15}^{45} tf_{gt} \times N_{gt}^f$$

où tf_{gt} est le taux de fécondité des femmes d'âge g de l'année t – en supposant que seules les femmes de 15 à 45 sont en âge d'avoir des enfants.

L'exercice proposé ici a deux buts

- réaliser une projection de population pour la France simple mais qui serait réaliste ;
- mettre en œuvre, comme méthode alternative, une micro-simulation.

2. — Une macro-projection pour la France

Pour faire simple, nous allons agréger la pyramide des âges par tranches d'âge vicennal. Nous distinguons ainsi les femmes âgées de 0 à 19 ans, de 20 à 39 ans, etc. De même pour les hommes. Les données nécessaires sont disponibles sur le site Web de l'INSÉÉ ; on obtient, sur le champ France métropolitaine au premier janvier 2006, après consolidation par tranche d'âge vicennal

Tranche d'âge vicennal	Femmes	Hommes
1	7 402 599	7 745 707
2	8 111 360	8 156 021
3	8 562 984	8 292 729
4	5 391 230	4 516 143
5	1 917 236	948 675

Il naît plus de garçons que de filles ; les femmes vivent cependant beaucoup plus vieilles que les hommes ; la faible fécondité des 25 dernières années fait que les effectifs de la tranche vicennale 1 sont plus réduits que ceux de la tranche 2. Voir aussi les résultats de la première PROC freq.

On trouve aussi sur le site de l'INSÉÉ une table de mortalité construite à partir des données des années 2002-2004. Cette table donne notamment les taux de survie à différents âges, par sexe. On a ainsi, en %

Age	Femmes	Hommes
10	100	99
30	99	98
50	97	93
70	87	73

Ces taux sont inconditionnels ; ils ne donnent pas la probabilité de survie une année sachant que l'on a déjà survécu toutes les années précédentes.

Pour simplifier l'exercice, on va supposer que seules les femmes de la deuxième tranche vicennale sont fécondes ; on retient un taux de natalité de 2 enfants par femme.

3. — Le programme SAS de la macro-projection

Le programme SAS tout d'abord

- charge la table pyramide (lignes 1 à 7) ;
- charge la table taux_survie (lignes 8 à 14).

Ensuite, la table des probabilités conditionnelles de survie est construite (lignes 15 à 32). Cette table est transposée pour pouvoir disposer de toutes les probabilités en une seule fois.

La macro-projection est programmée de la ligne 34 à la ligne 57. Notez l'utilisation explicite de l'instruction OUTPUT pour créer les observations. L'expression `tav = tav + 1` permet de construire l'observation pour la tranche d'âge vicennal suivante.

4. — Le programme SAS de la micro-simulation

L'enquête est engendrée de manière artificielle en utilisant les effectifs de la pyramide des âges initiale. C'est l'objet de l'étape DATA des lignes 68 à 74. La micro-simulation est programmée de la ligne 75 à la ligne 105. La fonction `ranuni` permet d'obtenir une réalisation d'une loi pseudo-aléatoire uniforme entre 0 et 1.

Le programme Sas

(1)	DATA pyramide ;
(2)	/* Source : Inséé, France métropolitaine, 1er janvier 2006. */
(3)	INPUT sexe \$ tav n @@; /* 'tav' : Tranche d'âge vicennal. */
(4)	CARDS ;
(5)	F 1 7402599 F 2 8111360 F 3 8562984 F 4 5391230 F 5 1917236
(6)	H 1 7745707 H 2 8156021 H 3 8292729 H 4 4516143 H 5 948675
(7)	RUN ;
(8)	DATA taux_survie ;
(9)	/* Source : Inséé, Table de mortalité des années 2002 - 2004. */
(10)	INPUT sexe \$ tav taux_survie @@; /* Le taux est en %. */
(11)	CARDS ;
(12)	F 1 99 F 2 99 F 3 97 F 4 87
(13)	H 1 99 H 2 98 H 3 93 H 4 73
(14)	RUN ;
(15)	DATA pcs_transp; /* Probabilités conditionnelles de survie (transposées). */
(16)	ARRAY pcsf pf1-pf4; RETAIN pf1-pf4; KEEP pf1-pf4; /* Les femmes. */
(17)	ARRAY pcsh ph1-ph4; RETAIN ph1-ph4; KEEP ph1-ph4; /* Les hommes. */
(18)	SET taux_survie END = FIN;
(19)	IF sexe EQ 'F' THEN pcsf{tav} = taux_survie ;
(20)	ELSE pcsh{tav} = taux_survie ;
(21)	IF FIN THEN DO ;
(22)	DO i = 1 TO 4 ;
(23)	pcsf{i} = pcsf{i} / 100 ;
(24)	pcsh{i} = pcsh{i} / 100 ;
(25)	END ;
(26)	DO i = 4 TO 2 BY -1 ;
(27)	pcsf{i} = pcsf{i} / pcsf{i-1} ;
(28)	pcsh{i} = pcsh{i} / pcsh{i-1} ;

Suite du programme sur la page suivante...

Le programme Sas (suite)

(29)	END ;
(30)	OUTPUT ;
(31)	END ;
(32)	RUN ;
(33)	PROC print DATA = pcs_transp; RUN ;
(34)	DATA pyramide_20 (KEEP = sexe tav n) ;
(35)	IF _n_ EQ 1 THEN SET pcs_transp ;
(36)	ARRAY pcsf pf1-pf4; ARRAY pcsh ph1-ph4 ;
(37)	SET pyramide ;
(38)	/* Les femmes en âge d'avoir des enfants. */
(39)	IF sexe EQ 'F' AND tav EQ 2 THEN DO ;
(40)	sauve_n = n ;
(41)	n = n * 2 / 2; /* 2 enfants par femme, une fille et un garçon. */
(42)	tav = 1; /* Tranche d'âge vicennal de l'enfant. */
(43)	sexe = 'H'; /* Les garçons. */
(44)	OUTPUT ;
(45)	sexe = 'F'; /* Les filles. */
(46)	OUTPUT ;
(47)	tav = 2; /* On restaure les valeurs des variables de la mère. */
(48)	n = sauve_n ;
(49)	END ;
(50)	IF tav NE 5 THEN DO ;
(51)	IF sexe EQ 'F' THEN proba = pcsf{tav} ;
(52)	ELSE proba = pcsh{tav} ;
(53)	n = n * proba ;
(54)	tav = tav + 1 ;
(55)	OUTPUT ;
(56)	END ;
(57)	RUN ;
(58)	PROC freq DATA = pyramide ;
(59)	TITLE 'Pyramide des âges - janvier 2006' ;
(60)	WEIGHT n ;
(61)	TABLE sexe*tav ;
(62)	RUN ;
(63)	PROC freq DATA = pyramide_20 ;
(64)	TITLE 'Pyramide des âges projetée - janvier 2026' ;
(65)	WEIGHT n ;
(66)	TABLE sexe*tav ;
(67)	RUN ;
(68)	DATA individus (KEEP = sexe tav poids) ;
(69)	SET pyramide ;
(70)	poids = 1000 ;
(71)	DO i = 1 TO n/poids ;
(72)	OUTPUT ;

Suite du programme sur la page suivante...

Le programme SAs (suite)

(73)	END ;
(74)	RUN ;
(75)	DATA individus_20 (KEEP = sexe tav poids);
(76)	IF _n_ EQ 1 THEN SET pcs_transp;
(77)	ARRAY pcsf pf1-pf4; ARRAY pcsh ph1-ph4;
(78)	SET individus;
(79)	IF sexe EQ 'F' AND tav EQ 2 THEN DO ;
(80)	sauve_sexe = sexe;
(81)	tav = 1; /* Tranche d'âge vicennal de l'enfant. */
(82)	/* Une chance sur deux d'avoir une fille. */
(83)	IF ranuni(12345678) < .5 THEN sexe = 'F'; ELSE sexe = 'H';
(84)	OUTPUT; /* Le premier enfant. */
(85)	/* 20 % des femmes ont 1 enfant, 60 % 2 et 20 % 3 enfants. */
(86)	IF ranuni(12345678) < (.6+.2) THEN DO ;
(87)	IF ranuni(12345678) < .5 THEN sexe = 'F'; ELSE sexe = 'H';
(88)	OUTPUT; /* Le deuxième enfant. */
(89)	IF ranuni(12345678) < .2/(.6+.2) THEN DO ;
(90)	IF ranuni(12345678) < .5 THEN sexe = 'F'; ELSE sexe = 'H';
(91)	OUTPUT; /* Le troisième enfant. */
(92)	END ;
(93)	END ;
(94)	tav = 2; /* On restaure les valeurs des variables de la mère. */
(95)	sexe = sauve_sexe;
(96)	END ;
(97)	IF tav NE 5 THEN DO ;
(98)	IF sexe EQ 'F' THEN proba = pcsf{tav};
(99)	ELSE proba = pcsh{tav};
(100)	IF ranuni(12345678) < proba THEN DO ;
(101)	tav = tav + 1;
(102)	OUTPUT;
(103)	END ;
(104)	END ;
(105)	RUN ;
(106)	PROC freq DATA = individus ;
(107)	TITLE 'Population par sexe et par tranches vicennales - janvier 2006';
(108)	WEIGHT poids ;
(109)	TABLE sexe*tav ;
(110)	RUN ;
(111)	PROC freq DATA = individus_20 ;
(112)	TITLE 'Population simulée par sexe et par tranches vicennales - janvier 2026';
(113)	WEIGHT poids ;
(114)	TABLE sexe*tav ;
(115)	RUN ;