Bulletin de l'Association des démographes du Québec

Décomposition des indices démographiques : quelques considérations générales

Réjean Lachapelle

Volume 2, Number 1, Special, 1973

URI: https://id.erudit.org/iderudit/305740ar
DOI: https://doi.org/10.7202/305740ar

See table of contents

Publisher(s)

Association des démographes du Québec

ISSN

0380-1713 (print)
1925-3478 (digital)

Explore this journal

Cite this article

Lachapelle, R. (1973). Décomposition des indices démographiques : quelques considérations générales. *Bulletin de l'Association des démographes du Québec*, *2*(1), 165–171. https://doi.org/10.7202/305740ar

DÉCOMPOSITION DES INDICES DÉMOGRAPHIQUES: QUELQUES CONSIDÉRATIONS GÉNÉRALES

Réjean Lachapelle,
Université de Sherbrooke
Faculté de médecine
Département des sciences du comportement

Les démographes font dans leurs travaux une grande consommation d'indices. Cela est si vrai qu'on pourrait, sans grande infidélité, écrire une histoire de la démographie qui se contenterait de retracer les discussions ayant pour thème le meilleur indice pour étudier ceci ou cela. Il y eut un engouement pour le taux intrinsèque d'accroissement naturel, les taux brut et net de reproduction, l'espérance de vie à la naissance, la standardisation directe...

Aujourd'hui encore on discute ferme pour savoir quel est le meilleur indice comparatif de mortalité (LIDDELL, 1960; SILCOCK, 1959; KILPATRICK, 1962; WOLFENDEN, 1962; SPIEGELMAN et MARKS, 1966; KITAGAWA, 1964, 1966; HILL, 1971). Plusieurs accordent leur faveur à la standardisation directe, d'autres penchent plutôt du côté de la standardisation indirecte; certains mettent en doute ces deux procédés bien connus et proposent d'autres méthodes de standardisation. Ce débat a au moins permis de faire justice d'un poncif qui figure encore dans la plupart des manuels d'introduction: si les données et les moyens de calcul (à l'époque de la règle à calcul on usait volontiers de cet argument, aujourd'hui il est manifestement anachronique) le permettent, employez la standardisation directe, sinon rabattez-vous sur l'indirecte. Je ne pense pas qu'on puisse dorénavant donner des consignes si simples aux étudiants. Il n'en demeure pas moins impérieux de proposer divers angles d'attaque pour faire face aux deux problèmes que tentaient de résoudre les standardisations directe et indirecte: d'une part, comment comparer l'intensité d'un phénomène dans deux ou plusieurs populations si peu de données sont disponibles (par exemple: comparaison de la mortalité de deux populations si pour l'une on ne dispose que du total des décès); d'autre part, comment dégager, pour le phénomène étudié, les différences essentielles entre diverses populations lorsque les données sont abondantes (par exemple) comparaison de la mortalité de plusieurs groupes d'individus si on dispose, pour chacun d'eux, de la composition par âge de la population et des décédés). Trois voies ont plus ou moins été explorées: 1o/ l'étude de la variance des divers indices (CHIANG, 1960, 1961, 1968; KEYFITZ, 1966, 1968); 2o/ l'analyse factorielle (LEDERMANN et BREAS, 1959; BOURGEOIS-PICHAT, 1963; LEBRAS, 1971, 1972); 3o/ la décomposition des indices démographiques (STOUFFER, 1951; KITAGAWA, 1955, 1964; CALOT, 1965). Dans la suite, on s'en tiendra à la troisième voie qui fut d'ailleurs la moins suivie des trois.

Décomposition des indices

Pourquoi utilise-t-on des indices? Pour répondre à cette question il n'y a pas de meilleur moyen, me semble-t-il, que d'examiner des travaux qui en font usage. Prenons, par exemple, l'étude de MM. Henripin et Légaré (1971) sur la fécondité canadienne. Pour décrire l'évolution de la fécondité de 1940 à nos jours on dispose d'une masse imposante d'informations; ces informations ne sont cependant pas toutes indépendantes les unes des autres: les redondances sont nombreuses. On souhaite donc rassembler dans quelques grandeurs choisies avec soin l'essentiel des renseignements disponibles. On aimerait en l'espèce pouvoir faire le départ, dans les variations du taux brut de natalité, entre les changements imputables aux modifications de la composition par âge et ceux dus aux variations des taux de fécondité par âge; mieux encore, on voudrait distinguer dans les variations imputables à la fécondité générale ce qui tient à la nuptialité de ce qui résulte des variations de la fécondité légitime et de la fécondité illégitime. Par ailleurs, il y aurait intérêt à traduire par un indice l'interaction entre la nuptialité et la fécondité illégitime.

Les auteurs de l'article ont proposé divers indices traduisant les variations du taux brut de natalité imputables à la composition par âge, aux taux de fécondité générale, à la nuptialité et aux taux de fécondité légitime. En fait, ils ont employé ce que nous appelons la décomposition multiplicative, mais sans en faire la théorie. Pour mieux tirer parti de ce procédé de décomposition, il nous apparaît important d'en faire une présentation systématique.

Décomposition multiplicative

Pour fixer les idées, imaginons que nous décomposons le taux brut de natalité.

t et t' représentent respectivement le taux brut de natalité de la population de référence et de la population étudiée.

t_i et t'_i représentent les taux de fécondité du groupe d'âges i.

a_i et a'_i représentent la fraction des femmes dans le groupes d'âges i.

Ces quantitiés vérifient les relations suivantes:

$$t = \sum_i a_i t_i$$

$$t' = \sum_i a'_i t'_i$$

$$\text{Soit} \quad I\,(A,\ F) = \frac{t'}{t} = \frac{\Sigma\, a'_i t'_i}{\Sigma\, a_i t_i} = \frac{\Sigma a_i t_i \left(\dfrac{a'_i t'_i}{a_i t_i} \right)}{\Sigma\, a_i t_i}$$

$$\text{Posons} \quad W_i = \frac{a_i t_i}{\Sigma a_i t_i} \quad \text{, d'où} \quad \Sigma\, W_i = 1$$

$$\text{Il vient:} \quad I\,(A,\ F) = \Sigma W_i \left(\frac{a'_i t'_i}{a_i t_i} \right) \qquad (1)$$

On souhaite faire apparaître les composantes de I (A, F), rapport de taux bruts de natalité. Soit respectivement I (A) et I (F) l'effet des changements de la composition par âge et l'effet des variations des taux de fécondité:

$$I\,(A) = \frac{\Sigma\, a'_i t_i}{\Sigma\, a_i t_i} = \frac{\Sigma\, a_i t_i \left(\dfrac{a'_i}{a_i} \right)}{\Sigma\, a_i t_i} = \Sigma\, W_i \left(\frac{a'_i}{a_i} \right) \qquad (2)$$

$$I\,(F) = \frac{\Sigma\, a_i t'_i}{\Sigma\, a_i t_i} = \frac{\Sigma\, a_i t_i \left(\dfrac{t'_i}{t_i} \right)}{\Sigma\, a_i t_i} = \Sigma\, W_i \left(\frac{t'_i}{t_i} \right) \qquad (3)$$

Définissons aussi des indices conditionnels:

$$I\left({}^{A}/_{F}\right) = \frac{I\,(A,F)}{I\,(F)} = \frac{\Sigma\, a_i' t_i'}{\Sigma\, a_i t_i'} = \frac{\Sigma\, a_i t_i' \left(\frac{a_i'}{a_i}\right)}{\Sigma\, a_i t_i'} \qquad (4)$$

variations imputables aux changements de la composition par âge, la fécondité étant fixée au niveau courant;

$$I\left({}^{F}/_{A}\right) = \frac{I\,(A,F)}{I\,(A)} = \frac{\Sigma\, a_i' t_i'}{\Sigma\, a_i' t_i} = \frac{\Sigma\, a_i' t_i \left(\frac{t_i'}{t_i}\right)}{\Sigma\, a_i' t_i} \qquad (5)$$

variations imputables aux changements des taux de fécondité, la composition par âge étant fixée au niveau courant.

Il y a *indépendance* entre les composantes A et F si:

$$I\left({}^{A}/_{F}\right) = \frac{I\,(A,F)}{I\,(F)} = I\,(A)$$

Dans ce cas particulier, on aura:

$$I\,(A,F) = I\,(A)\;I\,(F), \text{ c'est-à-dire:}$$

$$\Sigma\, W_i \left(\frac{a_i'}{a_i}\,\frac{t_i'}{t_i}\right) = \left[\Sigma\, W_i \left(\frac{a_i'}{a_i}\right)\right] \left[\Sigma\, W_i \left(\frac{t_i'}{t_i}\right)\right]$$

Posons: $$\begin{aligned} COV\,(I(A),\, I(F)) &= \Sigma\, W_i\, (\, I(a_i) - I(A))(I(t_i) - I(F)) \\ &= \Sigma\, W_i\; I(a_i)\; I(t_i) - I(A)I(F) \\ &= I\,(A,F) - I(A)\; I(F) \qquad (6) \end{aligned}$$

où $I\,(a_i) = \frac{a_i'}{a_i}$ et $I\,(t_i) = \frac{t_i'}{t_i}$

Si les variations de la fécondité et les modifications de la composition par âge s'effectuent sans corrélation linéaire, alors la *covariance* entre I (A) et I (F) est nulle et, de ce fait, il y a *indépendance* entre A et F. (Remarquons que le vocable *indépendance* n'a pas ici la même signification que dans la théorie des probabilités. Il serait sans doute préférable d'employer le mot multiplicativité).

On appelle *interaction* des composantes A et F le rapport:

$$\Gamma\ (A,F) = \frac{COV(I(A),I(F))}{I\ (A)\ I\ (F)} = \frac{I(F/A)-I(F)}{I\ (F)}$$

$$= \frac{I(A/F)-I(A)}{I\ (A)} \qquad (7)$$

L'interaction est nulle s'il y a indépendance; de plus, elle a toujours le même signe que la covariance.

A l'aide des grandeurs que nous venons de définir, procédons maintenant à la décomposition de I (A,F). En vertu de la relation (5), on peut écrire:

$$I\ (A,F) = I(A)\ I(F/A)$$

Or, par la relation (7),

$$I\ (F/A) = \Gamma(A,F)\ I\ (F) + I\ (F)$$

$$= I\ (F)\ (\ 1 + \Gamma(A,F))$$

Finalement:

$$I\ (A,F) = I\ (A)\ I\ (F)\ (\ 1 + \Gamma(A,F)) \qquad (8)$$

Cette équation permet de donner une interprétation de l'interaction: il s'agit de l'écart à la multiplicativité (indépendance) des composantes.

Par ailleurs, la relation (8) permet également d'écrire:

$$I\ (A,\ F) = \sqrt{(I\ (A))^2\ (1+\Gamma\ (A,F))}\sqrt{(I\ (F))^2\ (1+\Gamma\ (A,F))}$$

$$= \sqrt{I\ (A)\ I\ (A/F)}\ \sqrt{I\ (F)\ I\ (F/A)}$$

$$= \quad I^*\ (A)\ I^*\ (F) \qquad (9)$$

où $I^*\ (A) = \sqrt{I\ (A)\ I\ (A/F)}$ et $I^*\ (F) = \sqrt{I\ (F)\ I\ (F/A)}$

Si (A,F) est faible, on peut considérer que I* (A) [(resp. I* (F)] traduit l'effet de la composante A (resp. F).

Décomposition additive

On emploiera les mêmes symboles que pour la décomposition multiplicative.

Posons:

$$\Delta\ (A,F) = \frac{\Delta\ (t)}{t} = \frac{t'-t}{t} = \frac{\Sigma\ (t_i'a_i' - t_ia_i)}{\Sigma\ a_it_i}$$

$$= I\ (A,F) - 1 \qquad (10)$$

C'est cette différence relative que nous souhaitons décomposer.

La variation relative imputable à la composition par âge correspond à:

$$\Delta\ (A) = \frac{\Sigma\ t_i\ (a_i' - a_i)}{\Sigma\ a_it_i} = I\ (A) - 1 \qquad (11)$$

et pour F à:

$$\Delta\ (F) = \frac{\Sigma\ a_i\ (t_i' - t_i)}{\Sigma\ a_it_i} = I\ (F) - 1 \qquad (12)$$

On définit également des indices de variation conditionnelle:

$$\Delta\ (A/F) = \Delta\ (A,F) - \Delta(F)$$

$$= I\ (A,F) - I(F)$$

$$= \frac{\Sigma\ t_i'\ (a_i' - a_i)}{t} \qquad (13)$$

$$\Delta\ (F/A) = \Delta\ (A,F) - \Delta\ (A)$$

$$= I\ (A,F) - I\ (A)$$

$$= \frac{\Sigma\ a_i'\ (t_i' - t_i)}{t} \qquad (14)$$

Posons

$$\Delta\ (A \times F) = \Delta\ (A/F) - \Delta\ (A) \qquad (15)$$

On montre facilement que:

$$\Delta\ (A \times F) = \Delta\ (F/A) - \Delta\ (F)$$

En vertu de (13) et de (15):

$$\Delta\ (A,F) = \Delta\ (A) + \Delta\ (F) + \Delta\ (A \times F) \qquad (16)$$

Δ (A X F) représente un écart à l'additivité.

La décomposition additive peut de prime abord paraître artificielle. Il n'en est rien. Pour le montrer, il suffit de remplacer de manière appropriée le signe ' par le symbole de l'accroissement. En effet:

$$I\ (A,F) = 1 + \Delta\ (A,F) = \frac{\Sigma\ a_i't_i'}{\Sigma\ a_it_i} = \frac{\Sigma\ (a_i + \Delta a_i)\ (t_i + \Delta t_i)}{\Sigma a_it_i}$$

$$= \frac{\Sigma\ a_it_i}{\Sigma\ a_it_i} + \frac{\Sigma\ a_i\Delta t_i}{\Sigma\ a_it_i} + \frac{\Sigma\ t_i\Delta a_i}{\Sigma\ a_it_i} + \frac{\Sigma\ \Delta a_i\Delta t_i}{\Sigma\ a_it_i}$$

$$= 1 + \Delta\ (F) + \Delta\ (A) + \Delta\ (A \times F)$$

Exemple d'application La natalité du Québec de 1961 à 1970

Les procédés de décomposition que nous venons de décrire ont été utilisés pour analyser l'évolution du taux brut de natalité du Québec de 1961 à 1970. Les résultats apparaissent aux graphiques I et II[1].

Les courbes ont manifestement un air de parenté. Dans les deux cas, il est clair que les modifications de la composition par âge de 1961 à 1970 ont eu pour résultat d'atténuer partiellement les effets déprimants de la fécondité. Pour l'essentiel, les changements dans la composition par âge résultent de l'arrivée des femmes nées après 1945 aux âges de forte fécondité. Quant à l'explication de la baisse de la fécondité, elle déborde évidemment notre propos...

Il ne faut pas s'étonner que la valeur de l'interaction reste dans l'ensemble assez faible. En effet, on voit mal pourquoi les variations de la fécondité seraient liées à celles de la composition par âge. Dans une telle situation il semble naturel d'accorder préférence au procédé de décomposition dont les valeurs de l'interaction sont les plus faibles. Dans le cas étudié, ce critère conduit à adopter la décomposition multiplicative. En général, à cause de l'habitude de comparer des rapports plutôt que des différences, les démographes emploieront plus volontiers la décomposition multiplicative que la décomposition additive. Une telle pratique ne m'apparaît pas dangereuse, à moins que l'examen de l'interaction (ou des interactions dans le cas de décomposition à plus de deux composantes) favorise nettement la décomposition additive.

Résumé et conclusion

Souvent en démographie les informations sont tellement touffues qu'il faut se résoudre à en sacrifier pour y voir clair. Ce qui ne signifie nullement rejet d'une partie des données, mais plutôt agencement de celles-ci de manière à pouvoir dégager l'essentiel. Plusieurs procédés statistiques tentent de relever cette gageure. Nous avons présenté ici deux méthodes de décomposition des indices démographiques: la décomposition multiplicative et la décomposition additive. Elles ont l'avantage de s'étendre sans difficulté à plus de deux composantes et de se prêter aisément à des interprétations de nature démographique — en particulier la multiplicative. Toutefois, il n'en faut pas trop attendre, car, à la vérité, il ne s'agit que des méthodes de «dégrossissage». Sans contact intime avec les informations de base, on ne peut aller au-delà des généralités; mais, dépourvu de tels procédés de synthèse l'analyste se perd facilement dans l'épaisse jungle des données.

1. Ils sont tirés d'un travail fait en mars 1973 par un groupe d'étudiants (Lourdès FLOR MARTIN, René LECORRE, Madeleine ROCHON-LESAGE et Maria-Concepcion SEGOVIA-CUEVAS) dans le cadre du cours DEM 731 donné à l'Université de Montréal.

Bibliographie

BOURGEOIS-PICHAT, Jean, «Application of Factor Analysis to the Study of Mortality», *Milbank Memorial Fund*, 1963, pp. 194-229.

CALOT, Gérard. *Cours de statistique descriptive.* Dunod, 1965.

CHIANG, Chin Long, «A stochastic study of the life table and its applications: II. Sample variance of the observed expectation of life and other biometric functions», *Human Biology*, 1960, pp. 221-238.

CHIANG, Chin Long, «Standard error of the age-ajusted», *Vital Statistics, Special Reports Selected Studies* (National Center for Health Statistics), 1961, pp. 275-285.

CHIANG, Chin Long. *Introduction to Stochastic Processes in Biostatistics.* John Wiley, 1968.

HENRIPIN, Jacques et LEGARE, Jacques, «Recent Trends in Canadian Fertility», *Rev. canad, Soc. & Anth./ Canad. Rev. Soc. & Anth.*, 1971, pp. 106-118.

HILL, G.B., «The Use of Vital Statistics and Demographic Information in the Measurement of Health and Health Care needs», in *Methods of Health Care Evaluation*, publié sous la dir. de Sackett (D.L.) et BASKIN (M.S.), McMaster University, Hamilton (Canada), 1971, chap. 2.

KEYFITZ, Nathan, «Sampling Variance of Standardized Mortality Rates», *Human Biology*, 1966, no 3, pp. 309-317.

KEYFITZ, Nathan. *Introduction to the Mathematics of Population.* Addison-Wesley, 1968.

KILPATRICK, J., «Occupational Mortality Indices», *Population Studies*, nov. 1962, pp. 175-187.

KITAGAWA, Evelyn M., «Components of a difference between two rates», *Journal of the American Statistical Association*, 1955, pp. 1168-1194.

KITAGAWA, Evelyn M., «Standardized Comparisons in Population Research», Demography, 1964, pp. 296-315.

KITAGAWA, Evelyn M., «Theoretical Considerations in the Selection of a Mortality Index, and some Empirical Comparisons», *Human Biology*, 1966, no 3, pp. 293-308.

LEBRAS, Hervé, «Géographie de la fécondité française depuis 1921», *Population*, nov.-déc. 1971, pp. 1093-1124.

LEBRAS, Hervé, «La mortalité actuelle en Europe.I. Présentation et représentation des données», *Population*, 1972, pp. 271-293.

LEDERMANN, Sully et BREAS, Jean, «Les dimensions de la mortalité», *Population*, 1959, no 4, pp. 637-682.

LIDDELL, F.D.K., «The Measurement of Occupational Mortality», *British Journal of Industrial Medecine*, 1960, pp. 228-233.

SILCOCK, H., «The Comparison of Occupational Mortality Rates», *Population Studies*, nov. 1959, pp. 183-192.

SPIEGELMAN, Mortimer et MARKS, Hebert H., «Empirical Testing of Standards for the Age Adjustment of Death Rates by the Direct Method», *Human Biology*, 1966, no 3, pp. 280-292.

STOUFFER, Samuel A., «Standardization of Rates When Specific Rates are Unknown», in *Handbook of Statistical Methods for Demographers* de JAFFE (A.J.), 3 e éd., 1960, pp. 56-58.

WOLFENDEN, Hugh H., «On the Theoretical and Practical Considerations Underlying the Direct and Indirect Standardization of Death Rates», *Population Studies*, nov. 1962, pp. 188-190.