c. 2

11F0019F
STC
CA1
STC305
R26
no. 19

Statistique Canada Statistics Canada

Statistique Canada
Direction des études analytiques

Documents de recherche

Canada

BCTUT
8588959

M-121F
c. 2

RÉPARTITIONS DES RISQUES DE MORTALITÉ:
Une analyse de tables de mortalité

par

Geoff Rowe

n° 19

Division des études sociales et économiques
Statistique Canada
automne 1988

L'auteur assume seul la responsabilité des opinions formulées dans le présent document qui ne représentent pas nécessairement le point de vue de Statistique Canada.

RÉPARTITIONS DES RISQUES DE MORTALITÉ:

Une analyse de tables de mortalité

Geoff Rowe

Division des études sociales et économiques

Statistique Canada

Mots clés: Table de mortalité, hétérogénéité de la mortalité, méthode empirique de Bayes, estimation non paramétrique de densité.

AMS 1980: classifications par matière: primaire 92A15, secondaire 62G30

RÉSUMÉ

La présente étude examine la répartition des risques de mortalité au Canada pendant la période 1975-1977. La description repose sur des estimations non paramétriques des densités de probabilité des risques de mortalité dans des groupes d'âges quinquennaux. Ces estimations sont tirées de données sur la mortalité par âge dans des petites régions, après ajustement destiné à tenir compte de la variabilité d'échantillonnage.

La variabilité dans les données sur la mortalité représentait en majeure partie la variabilité existant dans les risques plutôt que l'instabilité des estimations en provenance de populations à faible effectif. La variabilité des risques est indiquée dans des tables de mortalité quantiles, lesquelles regroupent des tables de mortalité, dont chacune repose sur des estimations des risques à des quantiles fixes des répartitions des risques par âge.

Les estimations des quantiles d'espérance de vie à la naissance et d'âge médian au décès donnent des mesures sommaires de la répartition des risques. Selon les estimations, les intervalles interquartiles pour l'espérance de vie s'établissent à 3,2 ans chez les hommes et à 2,8 chez les femmes.

1. Introduction

La présente étude examine la description statistique de la variabilité, entre de petites régions, des quotients de mortalité par âge ${}_nq(x)$ dans l'intervalle d'âge $[x, x+n)$, en cas de survie jusqu'à l'âge x. Dans la plupart des cas, n est égal à cinq. Les données, qui sont des moyennes des statistiques annuelles portant sur la période 1975-1977, comprennent des estimations ${}_n\hat{q}(x)$ pour les deux sexes dans chacune des 259 régions géographiques du Canada. Les unités géographiques sont les divisions de recensement (DR), qui correspondent aux comtés ou à d'autres unités administratives infraprovinciales. La définition exacte diffère d'une province à l'autre.

Les 259 DR représentent toutes les régions du Canada, à l'exception du Yukon, des Territoires du Nord-Ouest et de la partie la plus septentrionale du Québec (Nouveau-Québec). Bien que les populations de nombreuses DR soient importantes du point de vue du statisticien, elles sont généralement inférieures à la taille que les démographes jugent d'habitude requise pour la construction de tables de mortalité. Les démographes admettent souvent en postulat que des populations de 100 000 habitants ou moins fournissent des estimations instables des quotients de mortalité.

L'étude présente deux caractéristiques que l'on ne rencontre pas souvent dans l'analyse des tables de mortalité. Premièrement, on suppose que les estimations de ${}_{*}q(x)$ sont soumises à une variabilité qui est analogue à la variabilité d'échantillonnage. Deuxièmement, les observations sont censées venir d'une population qui possède une répartition hétérogène des risques de mortalité. L'hétérogénéité se rapporte aux divergences dans les niveaux de risque et provient des différences qui existent entre les individus (par exemple, régime alimentaire, mode de vie, etc.).

Chaque DR est traitée comme un échantillon de la population canadienne, qui est choisi sans tenir compte des conditions de mortalité. Par conséquent, chaque ${}_{*}\hat{q}(x)$ est une moyenne calculée pour des personnes dont les risques de mortalité sont censés être hétérogènes. La variance d'échantillonnage de ${}_{*}\hat{q}(x)$ est déterminée par la variance des risques inobservés au sein des DR et par la taille de la population des DR. Il existe des différences marquées entre la taille des populations des DR. Selon le recensement de 1976, les chiffres de la population des DR examinées dans l'étude allaient de 1 500 habitants à 2,1 millions d'habitants. Ainsi, une description de la répartition des risques nécessite l'utilisation simultanée des estimations ${}_{*}q(x)$, qui ne sont pas fiables au même degré.

L'étude vise à décrire d'une manière plus complète les conditions de mortalité au sein des populations nationales. A l'intérieur de chaque intervalle d'âge, on calcule des estimations non paramétriques des fonctions de densité de probabilité des risques de mortalité. Ces estimations fournissent une méthode souple d'évaluation de l'inégalité existant dans les

risques de mortalité (par exemple, au moyen d'indices de dispersion basés sur des quantiles).

Comme résultat de l'estimation de densité, on construit des tables de mortalité du moment en reliant les estimations des risques par âge qui correspondent à un quantile donné, et on les appelle tables de mortalité quantiles. Cette approche diffère du groupement plus conventionnel des estimations des risques par âge correspondant à une population définie géographiquement. On calcule les tables de mortalité quantiles pour des groupes d'âges (par exemple, en construisant des tables de mortalité quantiles abrégées). Ces dernières servent à illustrer une meilleure description de la mortalité pour une population nationale hétérogène. Dans d'autres buts que celui de donner un exemple, il serait préférable d'utiliser des tables de mortalité quantiles par année d'âge.

Le présent article insiste sur les tables de mortalité quantiles. Pour justifier la décision de les baser sur des données relatives aux DR, il faut tenir compte des différences de fiabilité entre les estimations $\hat{q}(x)$. C'est-à-dire que les tables de mortalité devraient refléter l'hétérogénéité dans la mortalité, plutôt que l'instabilité des estimations portant sur des petites régions. Il faut en outre que la variabilité entre les moyennes relatives aux DR soit aléatoire, pour que l'on puisse valablement décrire la répartition des risques par une densité de probabilité. Les deux sections qui suivent immédiatement cette introduction abordent ces aspects du problème d'estimation. La section 2 décrit l'estimateur $q(x)$ et une approximation de sa variance. Dans la section 3, on prétend que le

groupement géographique des données peut servir de base pour estimer la répartition des risques dans l'ensemble de la population canadienne.

L'estimation de la répartition des risques se fait en deux étapes: (1) la section 4 décrit l'ajustement des observations afin de déterminer la variabilité d'échantillonnage par une méthode linéaire empirique de Bayes, et (2) on obtient, à la section 5, des estimations non paramétriques ajustées de la densité de probabilité des risques, selon l'âge. La section 6 donne certains résultats explicatifs pour deux groupes d'âges, chez les hommes. La section 7 fournit des résultats tirés de certaines tables de mortalité quantiles. Enfin, la section 8 présente des remarques en guise de conclusion.

2. Variabilité d'échantillonnage

Le fait d'attribuer une variance d'"échantillonnage" aux estimations des quotients de mortalité peut être sujet à controverse si l'enregistrement des décès est complet et le sous-dénombrement du recensement négligeable. Dans ce cas, les taux de mortalité peuvent être considérés comme des valeurs exactes. Cette opinion ne tient aucun compte des facteurs ne se répétant pas et essentiellement imprévisibles qui influencent la mortalité d'une population pendant une période donnée (par exemple, températures extrêmes, etc.) Brillinger (1986) présente une base précise pour décrire la "variabilité naturelle", et applique un modèle à la mortalité des Canadiennes pendant la période 1926-1982. Divers autres modèles sont présentés dans les commentaires de discussion qui suivent l'étude de Brillinger.

On représente de la manière suivante les valeurs observées ${}_n\hat{q}(x)j$ pour la DR j selon l'équation de Chiang (1972), (en se servant de la notation actuarielle):

$$ {}_n\hat{q}(x)j = {}_nD(x)j / \{ [nP(x)j + (n-{}_na(x))\ {}_nD(x)j] / n \} $$

$$ = {}_nD(x)j/K(x)j, \qquad (1) $$

dans laquelle ${}_nD(x)j$ correspond au nombre annuel moyen de décès pendant la période 1975-1977 chez les personnes âgées de x à x+n ans, ${}_nP(x)j$ représente la population du recensement de 1976, et ${}_na(x)$ désigne le nombre moyen d'années vécues dans l'intervalle d'âge par une personne décédée pendant la période observée. Le dénominateur qui figure dans l'équation 1 est une estimation de la population annuelle exposée au risque. Dans le reste de l'étude, on traitera le dénominateur K(x)j comme s'il s'agissait d'une valeur exacte.

Les nombres de décès par âge et DR sont tirés de totalisations spéciales fournies par la Division de la santé, de Statistique Canada. Les chiffres de la population proviennent de la base de données du recensement de 1976. Signalons que les chiffres de la population ont été arrondis de façon aléatoire à des multiples de 5, conformément aux exigences requises pour les totalisations du recensement destinées à l'usage public.

Pour les groupes d'âges (0-1 an) et (1-5) ans, ${}_na(x)$ était tiré des tables de mortalité masculine et féminine abrégées pour le Canada pendant la période 1975-1977 (Nagnur, 1986). Pour tous les autres groupes d'âges, on admettait en postulat que ${}_na(x)$ avait la valeur 2,5 (c'est-à-dire, en supposant une répartition uniforme de décès). Ce dernier procédé représente une première approximation. On dispose d'autres approximations plus

complexes (voir, par exemple, Keyfitz et Frauenthal, 1975), mais il est difficile d'évaluer l'opportunité de ces méthodes lorsqu'on les applique à de petites populations qui sont soumises à des migrations importantes. Par conséquent, on n'utilise dans la présente étude qu'une première approximation, c'est-à-dire une stratégie préconisée par Hoem and Jensen (1982) et d'autres auteurs.

Désormais, on omettra, pour simplifier les expressions, le n souscrit indiquant l'étendue de l'intervalle d'âge ainsi que la notation représentant la dépendance à l'égard de l'âge. Toute l'analyse qui suit s'applique d'une manière indépendante à chaque intervalle d'âge.

On peut présenter un modèle statistique de la variabilité de $\hat{q}j$ de la manière suivante. Supposons que Qij soit une variable aléatoire de Bernoulli représentant des risques possibles de mortalité pour l'individu i, dans la DR j au sein de l'intervalle d'âge approprié. Qij prend la valeur 1 (qui représente un décès) avec la probabilité qij, et la valeur 0 avec la probabilité 1-qij. Ainsi, qij représente un risque personnel de décès.

Le nombre de décès hypothétiques dans la DR j est donné par la somme (supérieure à i) des risques personnels. Le nombre de décès observés (par exemple, Dj) correspond à la somme des réalisations $\hat{q}ij$ (prenant la valeur 1 en cas de décès de la personne et 0 dans les autres cas) de la variable aléatoire Qij, de sorte que l'équation 1 fournit un estimateur du risque moyen. La variance réelle du nombre de décès observés (en supposant des risques indépendants) est donnée par l'équation suivante:

$$Var(Dj) = \sum qij\ (1 - qij), \qquad (2a)$$

que l'on peut récrire de la manière suivante:

$$Var(Dj) = Kj\ q.j\ (1-q.j) - \sum (qij - q.j)^2, \qquad (2b)$$

dans laquelle q.j représente la moyenne des risques personnels. L'équation 2b indique qu'une variance binomiale de Dj donne la variance maximale dans toutes les répartitions des risques personnels Kj avec la moyenne q.j. On peut améliorer l'exactitude de l'approximation binomiale en réduisant la largeur des intervalles d'âge, pourvu que Kj reste large. Le résultat dépend de la diminution de toutes les valeurs qij dans la limite. Signalons, toutefois, que l'hypothèse des risques indépendants est peut-être trop audacieuse.

Le reste de la présente étude traitera des étapes ultérieures d'analyse, en supposant que l'approximation binomiale soit adéquate. Ainsi, l'approximation de la variance d'échantillonnage de $\hat{q}j$ est $\hat{q}j(1-\hat{q}j)/Kj$. Il est à noter que les différences entre les valeurs Kj (par exemple, la taille de l'échantillon) entraînent des différences de fiabilité entre les valeurs $\hat{q}j$, et que l'on se base sur ce fait pour décrire la composante de la variance comme "variance d'échantillonnage".

3. Interprétation des différences de mortalité entre les DR

La construction de tables de mortalité régionales et la présentation de données sur la mortalité dans des atlas peuvent donner l'impression qu'il convient de traiter des populations distinctes géographiquement comme si elles étaient fermées aux migrations ou comme si, d'une manière ou d'une autre, les différences de mortalité dépendaient directement de la région. Toutefois, étant donné que les limites géographiques des DR sont déterminées

sans tenir compte des conditions de mortalité, les divergences entre les valeurs $\hat{q}j$ ne représentent pas des différences systématiques entre les populations des DR.

Même lorsque certains facteurs de risque de mortalité sont nettement liés à la géographie, il ne s'ensuit pas nécessairement que la géographie offre une base importante pour la stratification de la population exposée au risque (par exemple, en formant des sous-populations qui comportent des risques relativement homogènes). Dans la mesure où les facteurs de risque comprennent des habitudes ou modes de vie personnels, on peut s'attendre à ce qu'ils varient considérablement au sein des populations des DR.

A titre d'exemple, citons le cas du mélanome malin cutané. L'exposition répétée au soleil constitue un facteur de risque dans certains cas de cancer de la peau, et des facteurs purement géographiques comme l'altitude et le climat peuvent exercer une influence directe. Toutefois, dans le cas du mélanome, des influences prédominantes comprennent la sensibilité de l'individu et l'intermittence de l'exposition. Par conséquent, on pourrait représenter plus exactement des strates homogènes de risque en fonction de la profession (les personnes travaillant à l'intérieur étant exposées à un plus grand risque que celles qui travaillent à l'extérieur), de la sensibilité (peau blanche, tendance à brûler, etc.), et de la situation socio-économique (bronzage récréatif, vacances d'hiver, etc.), plutôt qu'en fonction seulement de la géographie (Elwood et autres, 1985).

Les caractéristiques des individus, dont l'appartenance à une population de DR plutôt qu'à une autre résulte d'un classement arbitraire, constituent une source de variabilité non systématique dans les estimations du risque moyen. Le classement peut correspondre à une répartition aléatoire de la population canadienne. Ainsi, la répartition des moyennes des DR reflétera la répartition des risques d'individus typiques au sein de la population nationale. C'est-à-dire qu'une densité de probabilité constitue une description adéquate de la répartition des moyennes des DR, étant donné que la source de variabilité comporte une interprétation des fréquences.

4. Ajustement linéaire empirique de Bayes destiné à corriger la variabilité d'échantillonnage

On peut utiliser les méthodes linéaires empiriques de Bayes (LEB) pour estimer des vecteurs de moyennes aléatoires. On peut représenter une forme simple de méthode LEB comme estimation d'un vecteur paramétrique aléatoire $\emptyset$ au moyen d'une régression sur un vecteur de données observées y (Robbins, 1983). L'objectif est de minimiser $E\{\emptyset(y))-\emptyset\}^2$, où $\emptyset(y)$ est une fonction linéaire d'y estimant $\emptyset$. Dans la présente application, les méthodes LEB consistent à tasser chaque qj vers le centre de la répartition. L'ampleur du rétrécissement est grande si la variance d'échantillonnage d'un $\hat{q}j$ donné est importante (c'est-à-dire en cas d'ajustement destiné à corriger la variabilité d'échantillonnage).

Le rétrécissement des estimations dans le sens de leur moyenne globale fournit un compromis raisonnable entre l'utilisation de la moyenne globale

pour représenter chaque DR (c'est-à-dire en supposant effectivement que la variance entre les DR est attribuable à la variabilité d'échantillonnage) ou l'acceptation de chaque estimation de DR sans modification (c'est-à-dire en supposant effectivement que les DR représentent des populations sans rapport entre elles).

On utilise une transformation de stabilisation de la variance afin que chaque $\hat{q}j$ ait une variance d'échantillonnage (asymptotique) indépendante de q.j. La stabilisation de la variance est donnée par la modification de la transformation de l'arc sinus d'Anscombe (1949):

$$\hat{t}j = 2 \text{ arc sinus } (\sqrt{(Dj + 3/8)/(Kj + 3/4)}). \qquad (3)$$

La modification d'Anscombe donne une stabilité accrue aux DR dans lesquelles K est petit, mais son effet est négligeable pour la plupart des DR examinées dans la présente application. Efron et Morris (1975) utilisent également une transformation de l'arc sinus appliquée à des données binomiales dans le cadre d'une méthode LEB.

La variabilité entre les moyennes des DR est représentée dans le modèle suivant:

$$\hat{t}j = t. + (tj - t.) + uj = t. + ej, \qquad (4)$$

où $\quad tj = 2 \text{ arc sinus } (\sqrt{q.j}).$

Signalons que $Var(ej) = v + 1/Kj$, où $1/Kj$ représente la variance binomiale (d'échantillonnage) d'uj compte tenu de la transformation de l'arc sinus et v représente la variance qui dépasse la variance d'échantillonnage (c'est-à-dire la variance entre les écarts des DR (tj-t.)). Mentionnons en outre que $Cov(\hat{t}j, tj)=v$.

Les estimations de t et de v sont obtenues itérativement. L'estimation $\hat{t}.$ emploie l'inverse de la variance estimée de chaque tj comme coefficient de pondération:

$$\hat{t}. = \sum (\hat{t}j\ wmj) \ / \sum wmj, \qquad (5)$$

où $wmj = 1/(\hat{v} + 1/Kj)$ avec sommation de m DR. Remarquons que si $\hat{v}$ était assez grand, les coefficients de pondération seraient presque égaux (c'est-à-dire à peu près indépendants de Kj).

Une estimation de v peut prendre la même forme qu'une estimation du maximum de vraisemblance en cas de normalité, avec un facteur de correction m/(m-1). On peut utiliser cette formule pratique de calcul:

$$\hat{v}i+1 = [m/(m-1)] \sum (\hat{t}j-\hat{t})\ wvj\ ^2 \ / \sum wvj, \qquad (6)$$

où $wvj = [1/(1 + 1/\hat{v}iKj)]$, avec $\hat{v}i+1$ représentant une estimation actuelle et $\hat{v}i$ l'estimation antérieure. Il est à noter que $E(\hat{v}) = v$, que les écarts $(\hat{t}j-\hat{t})$ soient normaux ou autrement. Si l'on donne une valeur de départ à $\hat{t}$, on peut résoudre l'équation 6 d'une manière itérative. On estime alors $\hat{t}$ à partir de l'équation 5 avec de nouveaux coefficients de pondération, et on répète le cycle jusqu'à la convergence.

Le tableau 1 présente certains résultats des calculs de la variance. Il indique des coefficients de variation (CV) et des pourcentages de la variance qui dépassent la variance d'échantillonnage (appelés pourcentages d'excédent) pour chaque groupe d'âges. Il s'agit de mesures sommaires qui fournissent, d'une part, une comparaison des écarts-types non attribuables à l'échantillonnage par rapport à la moyenne et, d'autre part, la variance d'excédentaire relativement à la variance d'échantillonnage. Les CV oscillent entre 8 % et 28 %, ce qui indique que les écarts par rapport à la

moyenne sont, d'une manière typique, importants relativement à la moyenne. A l'exception du groupe de 0-1 an, les valeurs du pourcentage d'excédent sont toutes supérieures à 70 % (ce qui indique que la majeure partie de la variabilité présentée dans les données est une variabilité non attribuable à l'échantillonnage). Ces résultats montrent clairement qu'il existe des différences importantes entre les niveaux de mortalité présentés dans les données relatives aux DR. De même, la moyenne globale ne constitue pas un indicateur uniformément adéquat des risques courus dans la population.

On peut tirer des estimations LEB types de risque (q') de l'équation suivante:

$$t'j = 2 \text{ arc sinus } (\sqrt{q'j}) = \hat{t} + \left[\hat{v} / (\hat{v} + 1/Kj)\right] ej, \qquad (7)$$

dans laquelle le facteur de rétrécissement appliqué à ej estime le rapport entre $Cov(\hat{t}j, tj)$ et $/Var(\hat{t}j)$ (c'est-à-dire d'une manière analogue à celle d'une pente d'une droite de régression). Si l'on poursuit l'analogie avec la régression, la variance résiduelle $Var(t'j-tj)$ peut s'exprimer par la formule $Var(tj) - Cov(\hat{t}j,tj)^2/Var(tj)$, et est estimée par $\left[\hat{v}/(\hat{v} + 1/Kj)\right]$ $(1/Kj)$.

Un certain nombre d'auteurs (notamment Louis, 1984, ainsi que Gaver et O'Muircheartaigh, 1987) se sont inquiétés de la possibilité que les estimateurs LEB types n'effectuent un rétrécissement trop poussé. Par exemple, l'observation selon laquelle la variance d'échantillon $\sum (t'j-\hat{t})^2/m-1$ est inférieure à $\hat{v}$ implique que la fonction de répartition empirique (FRE) de t'j n'approchera pas de la FRE de tj. Louis a proposé l'emploi d'un estimateur LEB restreint qui fournit une approximation de la FRE de tj, laquelle minimise l'écart par rapport aux estimations LEB types (c'est-à-dire en minimisant

$\sum wcj(t*j-t'j)^2)$:

$$t*j = 2 \text{ arc sinus } (\sqrt{q*j}) = [wcj/(wcj+d1)]\ t'j+[d2/(wcj+d1)], \qquad (8)$$

où wcj est une estimation de $1/Var(t'j-tj)$.

Les paramètres d_1 et d_2 sont des multiplicateurs de Lagrange choisis de façon à ce que la moyenne de l'échantillon $t*. = \sum t*j/m$ et la variance de l'échantillon $\sum (t*j-t*.)^2/m-1$ fassent pendant à $\hat{t}$ et $\hat{v}$, respectivement. Les valeurs de d_1 pour chaque sexe et groupe d'âges étaient obtenues au moyen d'un algorithme sécant. A la convergence, les valeurs approximatives de d_1 donnaient des écarts-types de l'échantillon qui différaient de $\hat{v}$ de moins de 0,00001 (c'est-à-dire de moins de 0,01 % d'erreur dans tous les cas). Après avoir déterminé les valeurs de d_1, on calculait le paramètre d_2 en réarrangeant l'équation 8 et on obtenait des moyennes de l'échantillon qui concordaient exactement avec $\hat{t}$ dans chaque cas.

Les estimations LEB restreintes t*j sont des estimations raisonnables des quantiles de risque dans la mesure où elles sont chacune proches de l'estimation LEB type correspondante et ont un rapport de concordance étroit avec la moyenne et la variance estimées de $\hat{t}j$. Ainsi, les estimations t*j servent de base à l'évaluation des fonctions de densité de probabilité du risque de mortalité après ajustement destiné à tenir compte des différences dans la fiabilité de $\hat{q}j$.

5. Estimations de la fonction de densité du risque

L'estimation non paramétrique d'une fonction de densité de probabilité du risque de mortalité vise à fournir une indication des caractéristiques

des répartitions dont il faudrait tenir compte dans un modèle paramétrique approprié. En ce sens, cette étape de l'analyse constitue une phase préliminaire. Toutefois, l'estimation de densité peut également donner des estimations de voisinage de pseudo-vraisemblance et des estimations d'échelle du risque de mortalité, et faciliter l'interpolation vers des quantiles de risque séparés par des espaces égaux.

On peut évaluer la qualité de l'ajustement d'un modèle de référence paramétrique simple de t* en considérant la linéarité ou l'absence de linéarité dans l'association entre les quantiles t* et les quantiles de référence prévus. Dans la présente application, les quantiles de référence (désignés ci-après par la lettre z) sont tirés d'une fonction de densité bêta. La densité bêta se présente sous la forme suivante:

$$f(z) = [z^{b-1} (1-z)^{c-1}] / \text{Bêta}\ (b-c), \qquad (9)$$

où Bêta (b,c) est la fonction bêta. Les densités bêta peuvent s'adapter à une grande variété de formes de répartition et se prêtent à la présentation de variables aléatoires limitées (t* est limitée par zéro et π). Les paramètres b et c sont choisis pour faire pendant à $\hat{t}/\pi$ et $\hat{v}/\pi^2$.

Silverman (1985) décrit un algorithme de lissage pour des fonctions spline cubiques, destiné à la régression non paramétrique, qui sera utilisé ici à la fois pour fournir des estimations ajustées de t* et des estimations ajustées des premières dérivées dt*/dz. L'algorithme minimise la fonction du vecteur:

$$(1-h) [t*-t*(z)]^T W [t*-t*(z)] + h \int t*''(z)^2\ dt*(z), \qquad (10)$$

où t* est un vecteur des quantiles, t*(z) représente un vecteur des valeurs ajustées (à partir de la régression sur z), t*''(z) représente les secondes

dérivées des valeurs ajustées, et W est une matrice diagonale des coefficients de pondération. Les dérivées $dt*(z)/dz$ estiment le rapport entre la probabilité (connue) de z et la probabilité de $t*(z)$. Ainsi, on peut représenter la non-linéarité dans l'association entre $t*$ et z par une densité empirique qui diffère de la densité bêta ajustée.

Le paramètre h figurant dans l'équation 10 est un paramètre de lissage, qui pondère par voie différentielle l'exactitude locale et l'ajustement global. A mesure que h s'approche de 1,0, l'algorithme représente la relation entre $t*$ et z comme une transformation d'échelle de localisation (c'est-à-dire que la linéarité implique que $t*''(z)=0$). A mesure que h s'approche de 0,0, l'algorithme comprend implicitement autant de paramètres qu'il y a d'observations (c'est-à-dire à condition d'avoir une interpolation dans laquelle $t*-t*(z)=0$). On peut obtenir une valeur appropriée de h en minimisant les erreurs de contrevalidation (Silverman, 1984).

Comme les valeurs ajustées sont des quantiles estimés, on peut calculer W à partir de l'inverse des estimations des variances $t*(z)$:

$$Var\{t*(z)\} = s^2 \left[(p(1-p)) / (m\ f\ t*(z)\{^2)\right], \qquad (11)$$

où p est une estimation appropriée de la FRE, m représente le nombre d'observations, et s^2 est un scalaire (Parzen, 1979). On obtient les ordonnées de probabilité $f\{t*(z)\}$ à partir des ordonnées de probabilité de z divisées par l'estimation correspondante de $dt*(z)/dz$. Ainsi, on peut calculer par itération l'algorithme de lissage pour des fonctions spline en mettant W chaque fois à jour grâce aux estimations améliorées de $f\{t*(z)\}$ et s^2. La minimisation des erreurs de contrevalidation et l'usage itératif de

l'équation 11 ont permis d'obtenir des valeurs de h (pour chaque groupe d'âges), qui étaient situées dans l'intervalle de 0,76 à 0,99.

On trouve dans les études de Bofinger (1975) et Parzen (1979) des détails sur l'usage analogue de quantiles et de quantiles modèles en vue d'obtenir une estimation non paramétrique de densité. On justifie l'utilisation du lissage pour une fonction spline en considérant cette dernière comme une fonction de lissage (noyau) des moyennes mobiles, dotée d'une largeur de bande locale proportionnée à la variance locale (Silverman, 1984).

On se sert ensuite des estimations $t^*(z)$ pour construire des tables de mortalité. A cette fin, on calcule $\bar{q}^*$ en invertissant la transformation de l'arc sinus (c'est-à-dire $\bar{q}^*=[\sin(t^*(z)/2)]^2$) et on obtient des estimations de la fonction de densité de probabilité de $\bar{q}^*$ en appliquant le déterminant jacobien de la transformée inverse à $f(t^*(z))$.

6. Résultats explicatifs

Cette section présente des résultats intermédiaires qui illustrent l'analyse. La figure 1 donne un tracé de $\hat{q}$ par rapport à $\bar{q}^*$ pour les hommes faisant partie des groupes de 0-1 an et 50-54 ans. On a choisi ces groupes d'âges parce que les répartitions étaient bien distinctes, mais pas au point de masquer complètement les détails (par exemple en réduisant la résolution du tracé). La figure 2 présente des tracés correspondants des ratios de densité empirique. Les ratios de densité, qui sont estimés à partir des premières dérivées ajustées obtenues ci-dessus, soulignent les différences

qui peuvent exister entre les fonctions de densité de risque de mortalité empirique et les transformations des fonctions de densité bêta ajustée. En ce sens, les ratios de densité facilitent l'évaluation de la qualité de l'ajustement. Signalons que les ratios de densité représentés graphiquement ont été ajustés pour tenir compte du facteur d'échelle constant π .

La pente d'une ligne résistante ajustant $\hat{q}j$ comme fonction de $\bar{q}*j$ peut fournir un indice général des effets du rétrécissement et du lissage. Chez les hommes de 0-1 an, la pente est de 1,33, ce qui indique que la dispersion des risques de mortalité non ajustés est généralement plus élevée dans une proportion de 33 % que celle des risques ajustés. La pente correspondante pour les hommes de 50-54 ans s'établit également à 1,08. Chez les hommes, la pente maximale est de 1,33 (groupe de 0-1 an), alors qu'elle est de 1,21 chez les femmes (de 35-39 ans). Les valeurs minimales correspondantes s'établissent à 1,03 chez les hommes de 20-24 ans et à 1,06 chez les femmes de 65-69 ans. L'effet des ajustements sur la dispersion des estimations des risques de mortalité n'est pas considérable; cette observation n'est pas compatible avec l'hypothèse selon laquelle les estimations de la mortalité dans les petites régions sont généralement instables.

La figure 1 illustre les effets du rétrécissement et du lissage sur la fonction $\hat{q}j$. Dans le groupe de 0-1 an, l'extrémité supérieure (y compris sept valeurs aberrantes possibles ou plus) a été considérablement réduite. Signalons que ce groupe d'âges à la plus faible proportion de variance non attribuable à l'échantillonnage de tous les groupes d'âges, ainsi que l'indiquent les valeurs du pourcentage d'excédent présentées au tableau 1.

Par conséquent, ce groupe d'âges est plus fortement influencé par le rétrécissement que les autres groupes d'âges, comme le montre la pente de la ligne résistante mentionnée plus haut.

La figure 2 donne une indication des différences qualitatives qui peuvent exister entre la répartition réelle des risques de mortalité et la densité bêta qui était choisie comme répartition de référence. Par exemple, les ratios de densité pour les hommes de 50-54 ans s'établissent, en général, à près de 1,0. Ainsi, pour ce groupe d'âges, la densité bêta peut fournir un modèle adéquat de probabilité. Peut-être est-il plus opportun d'utiliser un modèle à deux paramètres (c'est-à-dire la moyenne et la variance), qui semble convenir pour décrire la répartition des risques de mortalité pour ce groupe d'âges. Par contraste, les ratios de densité pour le groupe de 0-1 an représentent, chez les enfants en bas âge, des risques de mortalité qui sont moins obliques et ont des extrémités plus courtes que ce serait le cas dans une densité bêta. Cela est indiqué par la forme de la fonction du ratio de densité, c'est-à-dire par le sommet au centre (valeurs $> 1,0$) et les valeurs peu élevées (valeurs $< 1,0$) aux extrémités. On trouve plus souvent dans d'autres groupes d'âges des ratios de densité semblables à ceux des hommes de 0-1 an (c'est-à-dire qui indiquent un manque d'ajustement) que des ratios de densité qui ressemblent à ceux des hommes de 50-54 ans. Cela impliquerait qu'il serait peut-être nécessaire d'utiliser trois ou quatre paramètres pour évaluer la répartition des risques de mortalité pour la plupart des groupes d'âges.

Les résultats présentés ici sont influencés par les paramètres de rétrécissement qui figurent dans l'équation 8 (c'est-à-dire d_1 et d_2) ainsi

que par le choix du paramètre de lissage qui figure dans l'équation 10 (c'est-à-dire h). Ainsi, bien que les résultats puissent indiquer qu'un modèle à quatre paramètres serait plus approprié qu'un modèle à deux paramètres, ils ne constituent pas une base pouvant servir à des tests de signification (par exemple, des tests du rapport des vraisemblances).

7. Tables de mortalité quantiles

Les tables de mortalité du moment représentent les conditions de mortalité du point de vue des durées de vie synthétiques achevées (par exemple, en formant des courbes de survie à partir de l'intégrale du produit des probabilités de survie par âge $(1-\hat{q})$). Les tables de mortalité quantiles constituent une généralisation des tables de mortalité du moment construites sur des sections qui traversent les répartitions ajustées de mortalité par âge de $\bar{q}*$. Ces sections relient les quantiles correspondants de divers groupes d'âges. Les courbes de survie construites peuvent servir à évaluer la variabilité pendant la durée de la vie (par exemple, l'espérance de vie à la naissance, ou l'âge médian au décès).

Le tableau 2 présente une comparaison de quotients de mortalité publiés antérieurement (Nagnur, 1986 - appelés estimations GLOBALES dans le tableau 2), avec des estimations qui utilisent les quotients de mortalité ajustés pour les DR et les fonctions de densité empirique. Cette comparaison vise à valider les résultats des tables de mortalité présentées ici en démontrant que, sur le plan de l'importance, les estimations sont comparables à des estimations plus conventionnelles. En outre, pour la plupart des groupes d'âges, l'ordre des valeurs MOYENNES, MÉDIANES et

MODALES indique clairement l'existence d'une asymétrie positive dans les répartitions de risque.

Toutefois, aucune des estimations présentées au tableau 2 n'est exactement comparable aux autres en tant qu'estimation de la même mesure de localisation. Plus précisément, les estimations GLOBALES et MOYENNES diffèrent tellement que les premières peuvent être présentées comme une moyenne pondérée de $\hat{q}j$ avec Kj comme coefficient de pondération, alors que les secondes se rapportent à une moyenne pondérée ayant pour coefficients de pondération les inverses des variances empiriques. L'utilisation d'une méthode de pondération semblable à celle employée dans l'équation 5 pourrait produire des différences nettement plus petites entre les coefficients de pondération par rapport aux différences existant entre les tailles de population. Ainsi, l'observation selon laquelle des estimations qui utilisent des données relatives aux DR sont souvent plus élevées que les estimations GLOBALES plus conventionnelles peut donner à entendre qu'il existe un biais dans les estimations GLOBALES. En raison des différences existant dans ces méthodes particulières d'estimation, la question de savoir quel est l'estimateur pouvant produire le plus grand biais est trop complexe pour qu'on puisse la résoudre ici. Néanmoins, des données de pondération selon la taille de la population risquent de produire des résultats inadéquats.

Le tableau 3 présente des estimations de la durée moyenne et médiane de vie tirées de certaines tables de mortalité quantiles. Ces estimations indiquent que la variabilité dans les risques courus au sein de la population canadienne est compatible avec une vaste gamme de durées de vie.

On peut obtenir une vérification partielle de ces résultats en comparant les estimations de l'extrémité inférieure avec un groupe de la population canadienne qui a notoirement connu une forte mortalité pendant la même période. En 1976, l'espérance de vie de la population des Indiens inscrits (représentant environ 1 % de la population canadienne) était de 60 ans pour les hommes et de 66 ans pour les femmes (Rowe et Norris, 1985). Les estimations basées sur les tables de mortalité quantiles sont moins extrêmes que les estimations relatives à la population des Indiens inscrits, ce qui peut indiquer que les estimations quantiles sont prudentes.

Le tableau 3 donne un aperçu des tendances récentes de la mortalité au Canada. Pendant les deux dernières décennies, l'espérance de vie au Canada s'est accrue d'au moins un an pendant chaque période de cinq ans. De telles augmentations ne sont pas importantes comparativement aux intervalles quantiles présentés au tableau 3. A ce propos, il est bon de se demander si ces tendances résultent d'améliorations générales de la santé ou d'améliorations qui sont particulières à une partie de la population, c'est-à-dire de se poser la question de savoir si les répartitions des risques relatives à des périodes antérieures seraient décalées vers le haut ou si leurs extrémités supérieures seraient allongées par rapport aux répartitions de 1976. La distinction influe sur nos prévisions des tendances futures de la mortalité.

Les tables de mortalité quantiles servent à concentrer l'attention sur deux éléments de la description des conditions de mortalité dans des populations hétérogènes. Il s'agit du besoin de tenir compte des différences entre individus du même âge, et de la nécessité de considérer

les différences qui peuvent exister dans le même groupe d'individus à divers âges. On peut construire des courbes de survie en déterminant à la fois une répartition des risques à chaque âge et en traçant une voie à travers la suite des répartitions de risques. Il y a lieu de croire que les fumeurs qui arrêtent de fumer suivent une voie différente de celle des fumeurs qui ne cessent pas de fumer. Les tables de mortalité quantiles représentent des parcours traversant les répartitions des risques, qui sont, en un sens, déterminés à la naissance.

8. Discussion

On peut résumer les principaux résultats de cette étude de la manière suivante:

(a) La variabilité entre les quotients de mortalité par âge, pour de petites régions, reflète en majeure partie la variabilité qui existe dans les risques de mortalité plutôt que l'instabilité des estimations relatives aux petites régions.

(b) Les différences entres les moyennes des DR sont suffisamment grandes pour avoir de l'importance sur les plans descriptif ou de l'élaboration des politiques, même après l'ajustement destiné à tenir compte de la variance d'échantillonnage et après le lissage. Par exemple, les tendances récentes de la mortalité nationale représentées par des changements dans l'espérance de vie peuvent refléter une contraction dans les extrémités des répartitions des risques plutôt que des décalages au centre.

(c) Les tables de mortalité quantiles extrêmes donnent une description directe de l'étendue des risques de mortalité dans une population hétérogène

en représentant les meilleures et les pires durées de vie synthétiques qui sont compatibles avec les données. Les tables de mortalité quantiles fournissent également des estimations de la dispersion près du centre des répartitions des risques de mortalité.

L'interprétation de ces résultats dépend de l'hypothèse selon laquelle les populations des DR peuvent constituer des échantillons représentatifs de personnes exposées à des risques indépendants. Cette opinion est très différente de celle qui suppose explicitement ou implicitement la présence d'effets particuliers à la région. Par ailleurs, c'est également une opinion qui s'accorde mieux avec des conditions de mortalité dans lesquelles la maladie contagieuse joue un rôle relativement secondaire alors que le mode de vie joue un rôle important.

Ces résultats pourraient inciter le chercheur à effectuer une analyse plus approfondie des données sur la mortalité fragmentées à un niveau de précision détaillé. Les études complémentaires pourraient se servir de variables de fragmentation qui ne peuvent être considérées comme non systématiques (comme l'état matrimonial), en plus d'utiliser un niveau de résolution géographique plus poussé . Mais, pour l'instant la répétition de ce travail sur plusieurs périodes pourrait constituer le meilleur complément de la présente analyse.

REMERCIEMENTS

Je remercie mes collègues Jane F. Gentleman et David Binder, dont les commentaires ont permis d'améliorer tant le contenu technique que la présentation de ce document. Je tiens également à remercier deux collaborateurs anonymes pour leurs commentaires.

TABLEAU 1

Mesures sommaires de la variabilité
dans les risques de mortalité de la population
des divisions de recensement
(L'arc sinus transforme la $\hat{t}j$ de $\hat{q}j$)

	HOMMES		FEMMES	
GROUPE D'ÂGES	Coefficient de variation	Pourcentage d'excédent	Coefficient de variation	Pourcentage d'excédent
0-1	14.7	50.1	16.3	50.1
1-4	28.4	82.1	26.2	74.6
5-9	24.8	77.8	27.5	76.1
10-14	24.7	79.7	27.3	76.0
15-19	23.0	91.3	25.0	81.5
20-24	25.4	92.8	27.3	81.0
25-29	21.4	86.7	26.0	78.7
30-34	20.5	84.0	24.0	76.9
35-39	19.1	82.7	20.2	73.6
40-44	16.4	82.7	16.8	72.3
45-49	13.7	83.5	17.6	80.7
50-54	11.5	84.2	14.3	78.5
55-59	11.3	87.6	13.2	81.9
60-64	9.6	87.8	12.1	82.7
65-69	8.7	86.7	12.5	84.9
70-74	9.3	89.5	9.7	80.6
75-79	9.0	88.2	10.1	75.7
80-84	8.2	84.6	11.1	88.0
85-89	9.2	85.0	9.1	77.7

Coefficient de variation = $100\sqrt{\hat{v}} / \hat{t}$

Pourcentage d'excédent - Variance qui dépasse la variance d'échantillonnage en pourcentage de la variance totale = $100\ (m\ \hat{v}) / \sum (\hat{v} + 1/Kj)$

$\hat{t}$ - moyenne estimée, $\hat{v}$ - variance non attribuable à l'échantillonnage
Kj - taille de la population, m - 259 divisions de recensement

Tableau 2

Comparaison entre des estimations de localisation
pour des quotients de mortalité q

ÂGE	ESTIMATIONS GLOBALES	MOYENNES	MÉDIANES	MODALES
HOMMES				
0-1	0.0141	0.0149	0.0146	0.0143
1-4	0.0032	0.0042	0.0039	0.0034
5-9	0.0023	0.0033	0.0032	0.0029
10-14	0.0021	0.0030	0.0028	0.0025
15-19	0.0075	0.0092	0.0087	0.0077
20-24	0.0093	0.0117	0.0111	0.0099
25-29	0.0074	0.0092	0.0088	0.0081
30-34	0.0078	0.0098	0.0095	0.0089
35-39	0.0107	0.0123	0.0118	0.0110
40-44	0.0164	0.0183	0.0178	0.0169
45-49	0.0273	0.0284	0.0279	0.0270
50-54	0.0439	0.0442	0.0437	0.0428
55-59	0.0693	0.0694	0.0687	0.0675
60-64	0.1053	0.1050	0.1054	0.1068
65-69	0.1557	0.1507	0.1499	0.1486
70-74	0.2281	0.2224	0.2219	0.2224
75-79	0.3234	0.3179	0.3165	0.3174
80-84	0.4508	0.4467	0.4453	0.4459
85-89	0.5989	0.6214	0.6207	0.6131
FEMMES				
0-1	0.0114	0.0123	0.0120	0.0115
1-4	0.0025	0.0035	0.0033	0.0032
5-9	0.0015	0.0024	0.0023	0.0020
10-14	0.0013	0.0020	0.0018	0.0016
15-19	0.0026	0.0036	0.0034	0.0030
20-24	0.0027	0.0036	0.0034	0.0030
25-29	0.0028	0.0036	0.0034	0.0028
30-34	0.0038	0.0047	0.0044	0.0039
35-39	0.0059	0.0072	0.0070	0.0067
40-44	0.0091	0.0103	0.0101	0.0097
45-49	0.0152	0.0151	0.0147	0.0144
50-54	0.0218	0.0227	0.0225	0.0220
55-59	0.0339	0.0342	0.0338	0.0330
60-64	0.0512	0.0513	0.0511	0.0515
65-69	0.0800	0.0795	0.0794	0.0804
70-74	0.1267	0.1282	0.1279	0.1287
75-79	0.2010	0.2037	0.2024	0.2010
80-84	0.3222	0.3318	0.3324	0.3337
85-89	0.4866	0.4965	0.4919	0.4817

ESTIMATIONS GLOBALES - telles que publiées dans l'étude de Nagnur (1986)- basées sur les rapports entre le nombre total de décès et la population totale.

ESTIMATIONS MOYENNES, MÉDIANES, MODALES - telles que déterminées par les estimations empiriques de Bayes des risques de mortalité et les densités de probabilité empirique.

Tableau 3

Estimations quantiles de la durée de vie

Quantile	Espérance de vie à la naissance		Âge médian au décès	
	HOMMES	FEMMES	HOMMES	FEMMES
1%	63.4	71.4	68.9	77.0
5%	65.5	73.4	70.5	78.4
10%	66.5	74.3	71.2	79.1
25%	68.2	75.7	72.5	80.2
50%	69.8	77.1	73.8	81.1
75%	71.4	78.5	75.2	82.1
90%	72.8	79.7	76.4	83.0
95%	73.7	80.5	77.1	83.6
99%	75.3	82.0	78.4	84.8
ESTIMATIONS MOYENNES	69.6	77.0	73.7	81.0
ESTIMATIONS MODALES	70.2	77.4	74.0	81.2
ESTIMATIONS GLOBALES	70.26	77.70	73.8	81.3

Les estimations MOYENNES et MODALES sont tirées des quotients de mortalité présentés au tableau 2.

ESTIMATIONS GLOBALES - l'espérance de vie présentée est celle qui est publiée dans l'étude de Nagnur (1986) - les âges médians au décès sont interpolés à partir de la courbe de survie publiée.

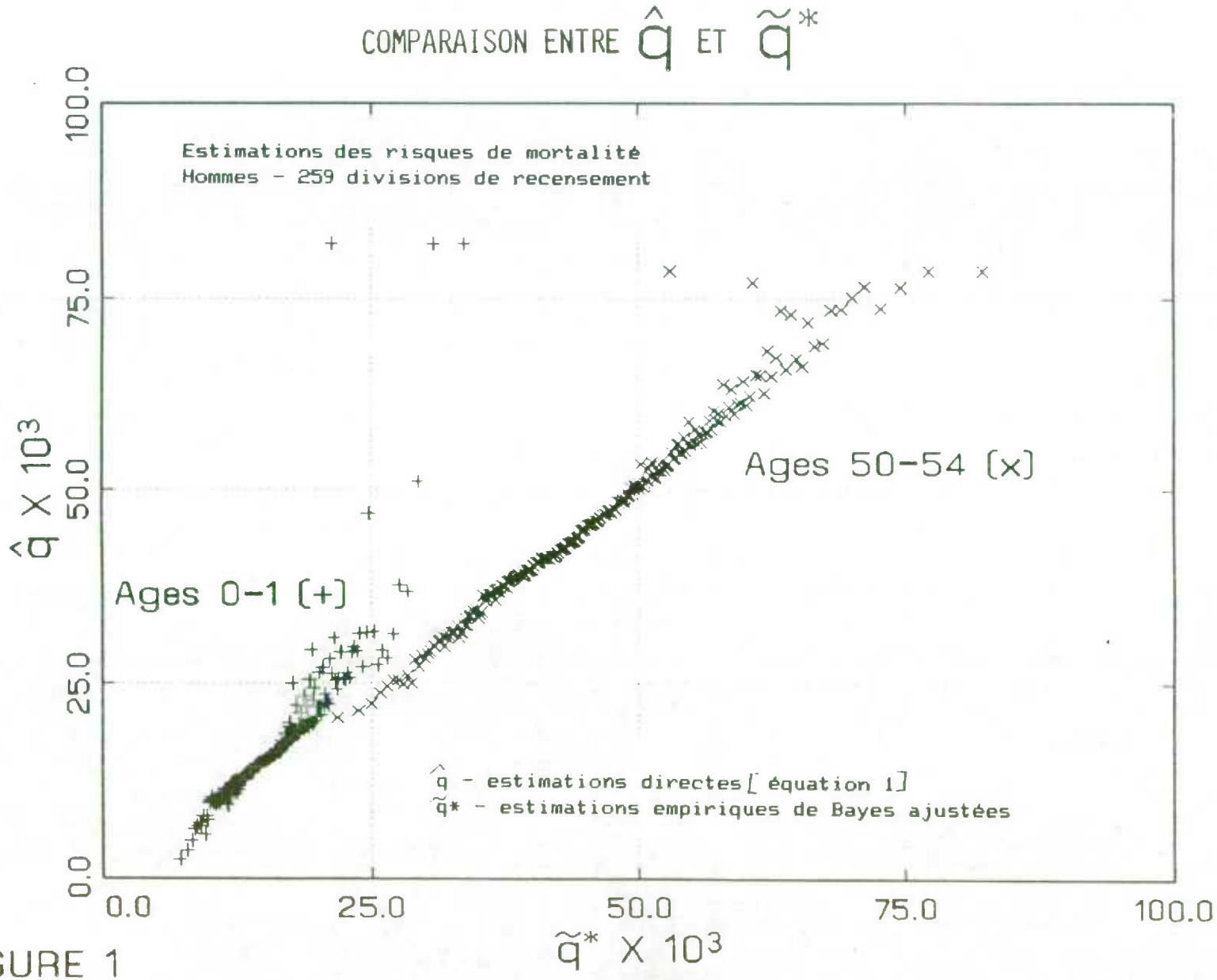
COMPARAISON ENTRE $\hat{q}$ ET $\tilde{q}^*$
Estimations des risques de mortalité
Hommes - 259 divisions de recensement
Ages 50-54 (x)
Ages 0-1 (+)
$\hat{q}$ - estimations directes [équation 1]
$\tilde{q}^*$ - estimations empiriques de Bayes ajustées
$\hat{q} \times 10^3$
$\tilde{q}^* \times 10^3$
0.0
25.0
50.0
75.0
100.0

FIGURE 1

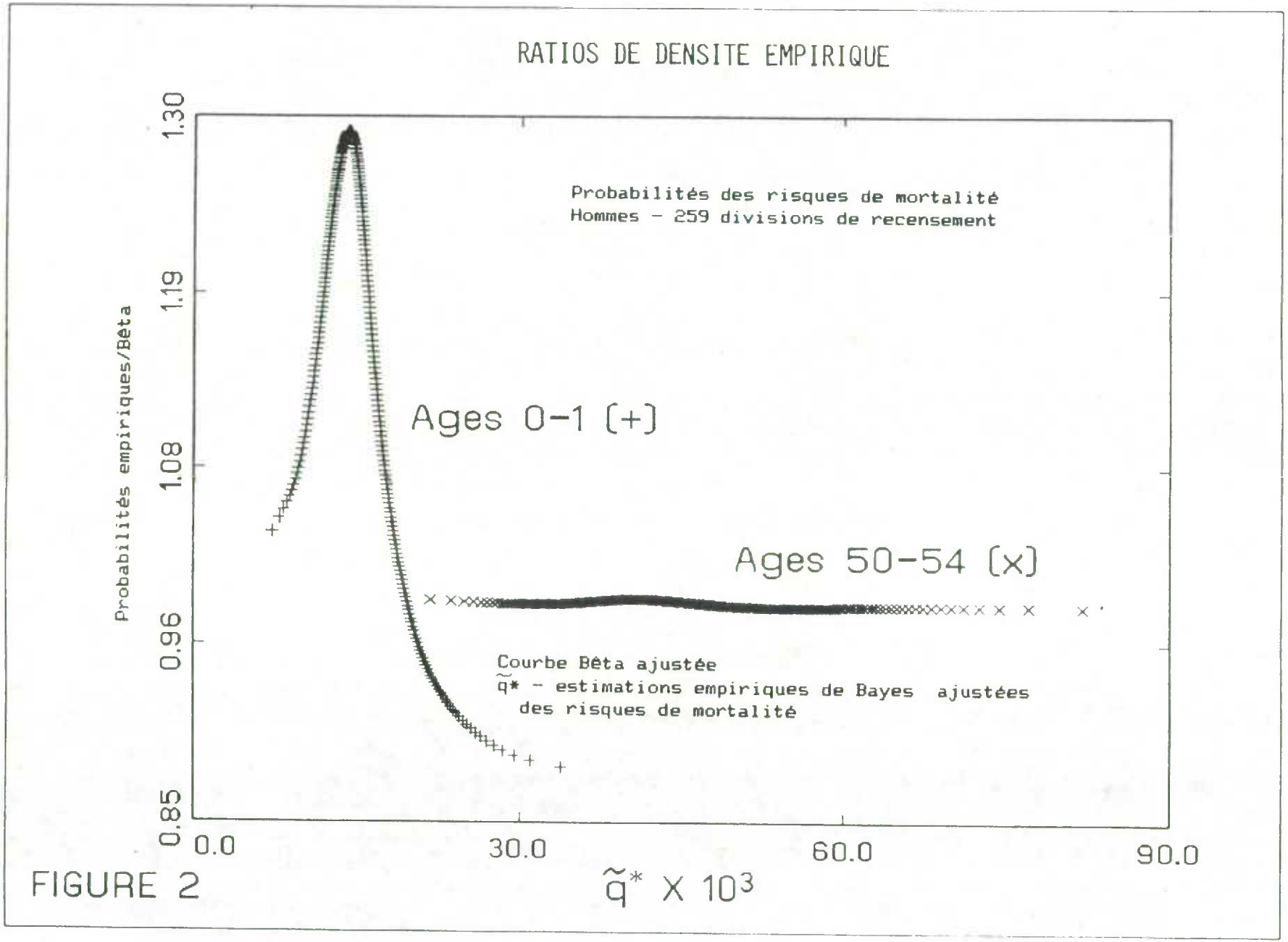
RATIOS DE DENSITE EMPIRIQUE
Probabilités des risques de mortalité
Hommes - 259 divisions de recensement
Ages 0-1 (+)
Ages 50-54 (x)
Courbe Bêta ajustée
q̃* - estimations empiriques de Bayes ajustées
des risques de mortalité
Probabilités empiriques/Bêta
1.30
1.19
1.08
0.96
0.85
0.0
30.0
60.0
90.0
q̃* X 10³

FIGURE 2

Bibliographie

Anscombe, F.J. (1948), "The transformation of poisson, binomial, and negative binomial data", Biometrika 35, 246-54.

Bofinger, E. (1975), "Estimation of a density function using order statistics", Austral. J. Statistics 17, 1-7.

Brillinger, David R. (1986), "The Natural Variability of Vital Rates and Associated Statistics", Biometrics 42, 693-734.

Chiang, C.L. (1972), "On constructing current life tables", JASA 67, 538-541.

Efron, Bradley et Morris, Carl (1975), "Data Analysis Using Stein's Estimator and Its Generalizations", JASA 70, 311-319.

Elwood, J.M., Gallagher, R.P., Hill, G.B., et Pearson, J.C.G., (1985), "Cutaneous Melanoma in Relation to Intermittent and Constant Sun Exposure - The Western Canada Melanoma Study", Int. J. Cancer 35, 427-433.

Gaver, Donald P., et O'Muircheartaigh, I.G. (1987), "Robust Empirical Bayes Analyses of Event Rates", Technometrics 29, 1-15.

Hoem, Jan M. et Jensen, Ulla Funck (1982), "Multistate Life Table Methodology: A Probabilist Critique", from Multidimensional Mathematical Demography, Kenneth C. Land et Andrei Rogers éditeurs, Academic Press, New York.

Keyfitz, Nathan et Frauenthal, James (1975), "An Improved Life Table Method", Biometrics 31, 889-899.

Louis, Thomas A. (1984), "Estimating a Population of Parameter Values Using Bayes and Empirical Bayes Methods", JASA 79, 393-398.

Nagnur, Dhruva (1986), Longévité et tables de mortalité chronologiques (abrégées) 1921-81, Canada et les provinces, n° 89-506 au catalogue, Statistique Canada, Ottawa.

Parzen, Emanuel, (1979), "Nonparametric Statistical Data Modeling", JASA 74, 105-131.

Robbins, Herbert, (1983), "Some Thoughts on Empirical Bayes Estimation", Annals of Statistics 11, 713-723.

Rowe, Geoff et Norris, M.J., (1985), "Projections de la mortalité des Indiens inscrits, 1982 à 1996", projet de rapport pour la Direction générale de la recherche, Orientation générale, Ministère des Affaires indiennes et du Nord Canada, Ottawa.

Silverman, B.W. (1984), "Spline Smoothing: The Equivalent Variable Kernel Method", Annals of Statistics 12, 898-916.

Silverman, B.W. (1985), "Some Aspects of the Spline Smoothing Approach to Non-parametric Regression Curve Fitting", J.R. Statis. Soc. B, 47, 1-52.

DIRECTION DES ÉTUDES ANALYTIQUES

DOCUMENTS DE RECHERCHE

No

1. RÉACTION COMPORTEMENTALE DANS LE CONTEXTE D'UNE SIMULATION MICRO-ANALYTIQUE SOCIO-ÉCONOMIQUE par Lars Osberg

2. CHÔMAGE ET FORMATION par Garnett Picot

3. DES PENSIONS AUX PERSONNES AU FOYER ET LEUR RÉPARTITION SUR LA DURÉE DU CYCLE DE VIE par Michael Wolfson

4. LA MODÉLISATION DES PROFILS D'EMPLOI DES CANADIENS AU COURS DE LEUR EXISTENCE par Garnett Picot

5. PERTE D'UN EMPLOI ET ADAPTATION AU MARCHÉ DU TRAVAIL DANS L'ÉCONOMIE CANADIENNE par Garnett Picot et Ted Wannell

6. SYSTÈME DE STATISTIQUES RELATIVES À LA SANTÉ: PROPOSITION D'UN NOUVEAU CADRE THÉORIQUE VISANT L'INTÉGRATION DE DONNÉES RELATIVES À LA SANTÉ par Michael Wolfson

7. PROJET PILOTE DE RACCORDEMENT MICRO-MACRO POUR LE SECTEUR DES MÉNAGES AU CANADA par Hans Adler et Michael Wolfson

8. NOTES SUR LES GROUPEMENTS DE SOCIÉTÉ ET L'IMPÔT SUR LE REVENU AU CANADA par Michael Wolfson

9. L'EXPANSION DE LA CLASSE MOYENNE: DONNÉES CANADIENNES SUR LE DÉBAT SUR LA DÉQUALIFICATION par John Myles

10. LA MONTÉE DES CONGLOMÉRATS par Jorge Niosi

Statistics Canada Library
Bibliothèque Statistique Canada
1010069482

11. ANALYSE ÉNERGÉTIQUE DU COMMERCE EXTÉRIEUR CANADIEN: 1971 et 1976 par K.E. Hamilton

12. TAUX NETS ET BRUTS DE CONCENTRATION DES TERRES par Ray D. Bollman et Philip Ehrensaft

13. TABLE DE MORTALITÉ EN L'ABSENCE D'UNE CAUSE POUR LE CANADA (1921 À 1981): Une méthode d'analyse de transition épidémiologique par Dhruva Nagnur et Michael Nagrodski

14. DISTRIBUTION DE LA FRÉQUENCE D'OCCURENCE DES SOUS-SÉQUENCES DE NUCLÉOTIDES, D'APRÈS LEUR CAPACITÉ DE CHEVAUCHEMENT par Jane F. Gentleman et Ronald C. Mullin

15. IMMIGRATION ET LE CARACTÈRE ETHNOLINGUISTIQUE DU CANADA ET DU QUÉBEC par Réjean Lachapelle

16. INTÉGRATION DE LA FERME AU MARCHÉ EXTÉRIEUR ET TRAVAIL HORS FERME DES MEMBRES DES MÉNAGES AGRICOLES par Ray D. Bollman et Pamela Smith

17. LES SALAIRES ET LES EMPLOIS AU COURS DES ANNÉES 1980: Évolution des salaires des jeunes et déclin de la classe moyenne par J. Myles, G. Picot et T. Wannell

18. PROFIL DES EXPLOITANTS AGRICOLES DOTÉS D'UN ORDINATEUR par Ray. D. Bollman

19. RÉPARTITIONS DES RISQUES DE MORTALITÉ: Une analyse de tables de mortalité par Geoff Rowe

Pour de plus amples renseignements, s'adresser au Président, Comité d'études des publications, Direction des études analytiques, Édifice R.H. Coats, 24ième étage, Statistique Canada, Parc Tunney's, Ottawa, Ontario K1A 0T6.