Crime, Histoire & Sociétés / Crime, History & Societies

Vol. 5, n°2 | 2001
Varia

Les limites intrinsèques du calcul de taux d'homicide

À propos des nouveaux standards proposés par E. Monkkonen

Bruno Aubusson de Cavarlay

Édition électronique
URL : https://journals.openedition.org/chs/735
DOI : 10.4000/chs.735
ISSN : 1663-4837

Éditeur
Librairie Droz

Édition imprimée
Date de publication : 1 janvier 2001
Pagination : 27-32
ISBN : 2-600-00664-8
ISSN : 1422-0857

Référence électronique
Bruno Aubusson de Cavarlay, « Les limites intrinsèques du calcul de taux d'homicide », *Crime, Histoire & Sociétés / Crime, History & Societies* [En ligne], Vol. 5, n°2 | 2001, mis en ligne le 02 avril 2009, consulté le 22 mars 2022. URL : http://journals.openedition.org/chs/735 ; DOI : https://doi.org/10.4000/chs.735

© Droz

Les limites intrinsèques du calcul de taux d'homicide

À propos des nouveaux standards proposés par E. Monkkonen

Bruno Aubusson de Cavarlay[1]

En réaction au texte d'Eric Monkkonen sur l'approfondissement des principes méthodologiques de la mesure de l'homicide, ce bref article conduit à mettre en doute la possibilité de respecter tous les critères de représentativité et de fiabilité attendus d'un indicateur statistique tout en restant dans la perspective de la mesure de la violence par un taux global d'homicide.

In reaction to Erik Monkkonen's text on further development of methodological principles of homicide measurement, this short article challenges the possibility of respecting all of the criteria of representativity and reliability one expects of a statistical indicator while retaining the measurement of violence through a global homicide rate.

L'état des recherches historiques sur l'homicide représente un cas de standardisation du questionnement quantitatif en sciences sociales. Un consensus s'est formé progressivement pour aborder le traitement des données dans un cadre fixé, ce qui permet ensuite la discussion d'interprétations et de constructions dites théoriques. Tout l'intérêt de l'article d'Eric Monkkonen est de revisiter cette phase de la formalisation quantitative, même au risque d'une critique potentiellement dévastatrice.

La définition de l'homicide et de ses catégories d'analyse n'est abordée que brièvement, Eric Monkkonen évoquant un consensus à ce propos. Le cas de l'infanticide, qu'il convient de traiter à part pour éviter un biais, n'épuise pas le sujet. N'y aurait-il donc rien d'autre à revoir au sujet par exemple de la prise en compte ou non des tentatives d'homicide (ceci pouvant avoir un lien avec la très pertinente analyse sur les pratiques médicales et les armes utilisées), de la délimitation entre homicides volontaires et involontaires, de l'incertitude qui peut demeurer dans la caractérisation de certaines morts violentes (suicides, accidents aux causes suspectes...)?

---

[1] Bruno Aubusson de Cavarlay est statisticien et sociologue, chercheur au C.N.R.S. (Centre de Recherches Sociologiques sur le Droit et les Institutions Pénales). Il mène des recherches quantitatives sur le fonctionnement de la justice pénale, ceci conduisant en particulier à une réflexion sur l'utilisation des statistiques criminelles dans une perspective historique. Sur ce point, il a publié notamment « Can Criminal Statistics Still Be of Scientific Use ? The French Criminal Justice System, 1831-1980 », *Historical Methods*, Spring 1993, vol. 26, 2, pp. 69-84 et « De la statistique criminelle apparente à la statistique judiciaire cachée », *Déviance et Société*, 1998, 22, 2, pp. 155-180.

Quelles garanties de stabilité ou de contrôle de ces définitions dans le temps et dans l'espace peut-on apporter ?

Il y a un lien entre ces questions de définition et la nature des sources utilisées. L'article de R. Roth (ci-après) relève les conséquences possibles du choix des sources sur les catégories d'homicide incluses dans le champ de l'enquête. Il relève aussi que l'interprétation de l'évolution des taux d'homicide ne peut se contenter d'une mesure globale mêlant l'homicide « domestique » aux autres homicides, lesquels peuvent se diviser à leur tour selon que les protagonistes avaient ou non des liens antérieurs et selon leur contexte. Le consensus existant pour faire de l'homicide, en sa globalité, un indicateur de violence en un temps donné pour un espace géographique donné ne peut faire oublier la diversité de cet ensemble : les interprétations théoriques doivent pouvoir s'appliquer à chacune de ses parties, ce qui suppose d'abord que les tendances observées leur sont bien communes.

Il y a aussi probablement un lien entre les catégories d'homicides, les sources utilisées et la fiabilité de la mesure et Eric Monkkonen propose de résoudre les problèmes posés à ce niveau en recourant à la méthode de capture-recapture. Cette méthode est initialement adaptée à l'estimation de la taille d'une population « sauvage » et suppose la remise en liberté des individus d'un premier échantillon après marquage. La proportion d'individus marqués dans un second échantillon représentatif dépend alors de la taille de la population totale, d'où l'estimation de ce paramètre. La littérature insiste sur les conditions de cette méthode : il faut qu'il y ait indépendance des deux échantillons. Si les animaux du premier échantillon sont devenus méfiants et échappent plus facilement à la recapture, ou si, affaiblis, ils sont attrapés plus facilement, la méthode fournit une estimation biaisée. La dépendance des deux échantillons couvre aussi le cas où certains individus ont une probabilité nulle d'être capturés lors des deux prélèvements.

Si l'on transpose au cas des homicides en confrontant les listes obtenues à partir de registres judiciaires (ou policiers) et du dépouillement de la presse, la méthode ne peut fournir une estimation fiable que si l'apparition dans une source ne dépend pas de l'apparition dans l'autre. Ce qui est probablement une hypothèse bien forte. Eric Monkkonen prévient cette objection en avançant que dans ce cas, on obtient une estimation par défaut certes, mais meilleure qu'à partir d'une seule source, du moins peut-on comprendre ainsi son argumentation. Or la variation du degré de dépendance entre les deux sources peut entraîner, toutes choses égales par ailleurs, des variations dans l'estimation fournie par l'estimateur de capture-recapture. Il est aussi probable que certains cas d'homicide échappent totalement à l'enregistrement policier ou judiciaire et à la presse : la proportion par rapport aux cas estimés demeurera toujours inconnue sans nouvelle information. Enfin, la méthode de capture-recapture suppose que tous les cas enregistrés sont des cas vrais. Si l'une des sources enregistre à tort certains faits comme des homicides, alors l'erreur se propage à l'estimation.

Je n'ai pas eu la possibilité de lire la communication sur laquelle ces affirmations sont fondées (Eckberg, 1998), mais je doute que l'estimation obtenue lève tous les problèmes de sources et de définition. Cela ne signifie pas pour autant que l'exercice soit inutile : le repérage et la correction apportée en cas de lacunes de la source judiciaire pour certaines années en est la preuve. En revanche, pour que cette méthode puisse servir d'épreuve à la thèse de Taylor ou à celle d'autres auteurs qui mettent en doute la fiabilité de la mesure policière ou judiciaire de l'homicide, il faudrait aussi

mettre à l'épreuve l'indépendance de la source judiciaire et de la source journalistique, ce qui paraît tout aussi difficile.

En posant la question de la pertinence des observations, Eric Monkkonen aborde conjointement ce qui pour le statisticien relève de la définition des unités de recueil des données, de leur représentativité et, éventuellement, de l'échantillonnage.

Les méthodes d'échantillonnage statistique sont conçues pour éviter de faire des choses aussi incongrues que d'évaluer le taux d'homicide aux USA à la fin du XX^{e} siècle sur la base des résultats de San Antonio et Gary. Comme il est (au moins à ce jour) difficile d'envisager que les historiens coordonnent leurs efforts pour travailler simultanément sur un « échantillon » de villes différentes pour une période donnée de façon à assurer une meilleure représentativité des résultats, on peut effectivement imaginer que le processus devienne cumulatif, chaque recherche apportant son observation supplémentaire à la courbe, avec les éléments de comparaison et d'homogénéisation indispensables. Comme dans tout échantillon géographique, les observations n'auront alors pas toutes le même poids, la même valeur informative.

Ces considérations vaudraient pour l'observation d'un autre type de crime. Pour l'homicide, il y a une particularité supplémentaire: il s'agit en général de très petits chiffres. Quelles que soient les précautions prises, la variabilité temporelle ou géographique du phénomène observé reste très grande. Comme on n'est pas encore en situation de produire des estimations fondées sur des échantillons représentatifs au sens statistique, cette variance n'intervient guère pour apprécier la fiabilité des conclusions. Et pourtant, la liste des 188 villes nord-américaines donnée en annexe du texte de Eric Monkkonen laisse imaginer l'ampleur du problème: en 1996 aux USA, la plage de variation du taux d'homicide est supérieure à celle des observations rapportées dans le graphique de Gurr si on excepte l'étrange point d'Oxford! Et encore, il ne s'agit que de villes de plus de 100 000 habitants!

Les résultats pour ces villes sont présentés en taux pour cent mille. Connaissant la population de référence, on peut se livrer à une analyse statistique de la variabilité du nombre et du taux d'homicides avec la taille des villes. Je n'entends pas le faire ici mais je relève qu'une évaluation rapide faite en divisant l'ensemble en trois groupes (par exemple moins de 150 000 habitants, de 150 000 à moins de 350 000, plus de 350 000 habitants) permet d'observer que la dispersion[2] des taux d'homicide décroît du premier groupe au troisième et que pour le premier groupe, la dispersion des taux d'homicide est encore en raison inverse de la taille. On peut aussi observer, cela va de pair, que le coefficient de corrélation entre le nombre brut d'homicides et la population est nettement plus faible pour le premier groupe que pour le troisième. Ceci aurait des conséquences pour la construction d'un échantillon de villes représentatif des 188 considérées si l'on devait procéder ainsi pour étudier quantitativement l'homicide. Opération délicate mais non impossible.

La situation pour les historiens est autrement plus difficile car les données mobilisées peuvent concerner des unités d'observation d'effectifs beaucoup plus faibles. Le commentaire d'Eric Monkkonen à propos du cas d'Oxford au XIV^{e} siècle pose bien le problème. Si dans une ville de 6 000 habitants, on enregistre pour une année

---

[2] Mesurée par exemple par le coefficient de variation, rapport de l'écart-type à la moyenne, soit 1,5 pour le premier groupe, 0,9 pour le second, 0,8 pour le troisième.

donnée un homicide, le taux est de 16,7 pour cent mille. Notons au passage que si l'unité de compte est la victime et si le meurtrier avait tué dans le même moment deux personnes, le taux aurait été de 33,3 pour cent mille. Ou alors un homicide peut à nouveau être enregistré les deux années suivantes, puis plus rien pendant vingt ans. Le taux est alors, sur ces vingt-trois ans de 2,2 en moyenne annuelle, valeur qui ne pourrait être observée pour aucune année en particulier.

La présentation que fait Eric Monkkonen de ce calcul permet de pousser plus loin l'interrogation. Il obtient le taux de 2,2 pour cent mille en rapportant les 3 homicides observés en 23 ans à une population de 138 000 habitants, obtenue en multipliant 6 000 par 23. On atteint par cette présentation le cœur de la mise en équivalence réalisée par le calcul de taux. La ville de 6 000 habitants imaginée serait équivalente à une ville de 138 000 habitants où l'on observerait en moyenne 3 homicides par an. Poussons cette présentation un peu plus loin encore. Si la première ville n'enregistre qu'un homicide au cours des 23 années suivantes en gardant la même population[3], si la seconde enregistre pendant une période 23 ans, un doublement de sa population et finalement un nombre moyen d'homicides de 2 par an, sera-t-on prêt à défendre qu'il s'agit d'une baisse équivalente de la violence dans les deux cas ?

L'une des façons de contrebalancer cette mise en équivalence trop brutale est de moduler le calcul d'un taux global par la prise en compte des critères discriminants relativement au phénomène étudié. Dès lors que l'âge est une variable importante pour l'analyse des homicides à un moment donné et que la structure par âge de la population concernée (dénominateur pour le calcul de taux) varie, il faut tenir compte des deux aspects dans l'évolution globale. Il y a ce qui résulte de l'évolution des taux d'homicide par classe d'âge et ce qui résulte de l'évolution de la structure d'âge de la population. La méthode appliquée en démographie, principalement pour étudier l'évolution de la mortalité, est celle dite des taux standards directs qui suppose le choix d'une structure de référence pour la population.

Le choix de cette structure de référence provient d'une convention lorsqu'il s'agit de mettre en harmonie des institutions publiant et analysant des données calculées en taux standards. On ne peut pas dire que cette décision est indifférente puisque les résultats obtenus en dépendent. L'article cité en référence à ce sujet par Eric Monkkonen (Anderson, Rosenberg, 1998) évoque clairement les conséquences de ce choix et les difficultés qu'il soulève. Le changement de référence aux USA (passage de taux standard 1940 à des taux standard 2000) entraîne une « réduction » du taux standard pour 1995 pour l'homicide (source causes de décès) de 10%. Certains résultats généraux concernant les taux de mortalité s'en trouvent effectivement modifiés (par exemple l'évolution de la mortalité par cancer ou la différence de mortalité entre Noirs et Blancs).

Le choix d'une population de référence est donc un élément très important. Il doit être justifié par rapport aux observations que l'on compare. On ne voit pas pourquoi la situation des USA en 2000 devrait servir de jauge pour l'ensemble des pays et des périodes où l'homicide est étudié.

En général, avec le calcul de taux standard, le sens des évolutions ne se trouve pas modifié mais seulement l'ampleur de ces évolutions. Anderson et Rosenberg

[3] Ou bien s'agissant d'une période ancienne, les bases manquant pour réviser l'estimation faite quant à la population.

exposent que ce résultat général se produit lorsque les taux spécifiques par âge ont le même profil d'évolution, à l'ampleur des variations près. Mais ils citent des exemples de cause de mortalité où ce n'est pas le cas et les données présentées montrent clairement que ce n'est pas le cas non plus pour l'homicide. D'après les statistiques sanitaires des USA, le taux d'homicide des moins de 25 ans augmente entre 1979 et 1995 (+23%) et ceux des 25-64 ans et des 65 ans et plus diminuent (respectivement –27% et –39%). Comme le passage de la référence de 1940 à 2000 pour la population des USA augmente la pondération des plus de 65 ans d'un facteur 4, ce changement de base conduit à afficher une baisse du taux standard d'homicide bien plus importante en base 2000 (-13,6% de 1979 à 1995) qu'en base 1940 (-7,8%). Un tel effet joue de même sur les chiffres proposés par Eric Monkkonen (source policière) pour la période 1980-1990 aux USA à New York (figure 4).

Si le renforcement des plus de 65 ans dans la population totale continue au même rythme, un historien américain pourra en 2040 reprendre cette courbe sur la base de taux standard 2040 et dire qu'en fait la période 1980-1990 a été une période de baisse très significative de l'homicide aux États-Unis, même sur la base des chiffres policiers.

C'est ce que suggère Eric Monkkonen après la standardisation des taux d'homicide sur la base 2000 pour les USA en soulignant que le résultat est l'inverse en termes d'évolution de celui que donnent les taux bruts pour certaines périodes. Ces situations montrent la limite de la méthode qui ne permet pas de résumer correctement l'évolution différentielle par classes d'âge ou de séparer un effet d'évolution générale et un effet de structure d'âge lorsque les évolutions par classe d'âge sont contradictoires. Seule une étude de l'évolution des taux par classe d'âge est alors pertinente.

Par ailleurs, le calcul de taux de mortalité standardisés selon l'âge vient de ce que la mort est un phénomène biologique certain dont la probabilité d'occurrence individuelle varie au cours de la vie. Parmi les causes de mortalité, l'homicide en est une assez particulière et elle présente un profil plutôt voisin des causes accidentelles que des causes biologiques. Autrement dit le profil par âge de l'homicide est plus lié à celui des auteurs, celui des victimes allant de pair, comme le souligne Eric Monkkonen.

Cette logique de la comparaison de population potentielle d'auteurs d'homicides à travers le temps et l'espace s'éloigne ainsi forcément d'une perspective démographique fondée sur les phénomènes biologiques de renouvellement ou de vieillissement. En général, des rapports entre mouvements démographiques et évolutions de la criminalité enregistrée ne sont observables qu'à l'occasion de changements brusques résultant de migrations ou d'accidents démographiques (liés à des guerres, des épidémies, ou de brusques variations de natalité). C'est aussi ce que suggèrent certains résultats présentés par Eric Monkkonen (migrations vers les villes au XIX^{e}) et sa remarque concernant la prise en compte de populations temporaires. De tout ceci, il convient très certainement de tenir compte dans la mise en perspective historique de long terme sachant que les points d'observation ou des portions de courbe d'évolution peuvent ou non refléter des situations démographiquement perturbées.

Une simple technique de standardisation de taux par âge mise au point par les démographes dans un tout autre but, a cependant l'inconvénient de noyer ces effets démographiques tout à fait significatifs dans un processus correctif qui valorise le point de vue contemporain d'une société où les classes âgées sont plus nombreuses que par le passé.

Toutes les recommandations d'Eric Monkkonen pour de nouvelles exigences méthodologiques sont fondées. Mais on peut se demander ce qu'il reste de la pertinence du taux d'homicide comme indicateur du niveau de violence sur de très longues périodes et sur des aires de taille et de population très variables. Ou au minimum si, au delà d'une indication toujours provisoire et révisable, une telle mesure peut fournir la matière première à des modélisations statistiques toujours plus complexes supposées valider des théories sur la genèse ou l'évolution de la violence.

Bruno Aubusson de Cavarlay
CESDIP-CNRS
43 boulevard Vauban
78280 Guyancourt
aubusson@ext.jussieu.fr

RÉFÉRENCES

Anderson, R. N., Rosenberg, H.M., Age standardization of death rates: implementation of the year 2000 standard, *National Vital Statistics Reports*, 1998, 47, n. 3, Available as http://www.cdc.gov/nchs/data/nvsr47_3.pdf.

Eckberg, D., *Stalking the elusive homicide: dual enumeration techniques and post-reconstruction South Carolina killings*, Paper presented at the Social Science History Association annual meeting, Chicago, 1998 and forthcoming *Social science history*, 2001.