|  |
| --- |
| Michel de Sève  Sociologue, professeur, département de sociologie, Université Laval  (1993)  “La mesure des effets d’inégalité et de discrimination avec des rapports purifiés.”  **LES CLASSIQUES DES SCIENCES SOCIALES** CHICOUTIMI, QUÉBEC <http://classiques.uqac.ca/> |



<http://classiques.uqac.ca/>

*Les Classiques des sciences sociales* est une bibliothèque numérique en libre accès, fondée au Cégep de Chicoutimi en 1993 et développée en partenariat avec l’Université du Québec à Chicoutimi (UQÀC) depuis 2000.



<http://bibliotheque.uqac.ca/>

En 2018, Les Classiques des sciences sociales fêteront leur 25e anniversaire de fondation. Une belle initiative citoyenne.

Politique d'utilisation  
de la bibliothèque des Classiques

Toute reproduction et rediffusion de nos fichiers est interdite, même avec la mention de leur provenance, sans l’autorisation formelle, écrite, du fondateur des Classiques des sciences sociales, Jean-Marie Tremblay, sociologue.

Les fichiers des Classiques des sciences sociales ne peuvent sans autorisation formelle:

- être hébergés (en fichier ou page web, en totalité ou en partie) sur un serveur autre que celui des Classiques.

- servir de base de travail à un autre fichier modifié ensuite par tout autre moyen (couleur, police, mise en page, extraits, support, etc...),

Les fichiers (.html, .doc, .pdf, .rtf, .jpg, .gif) disponibles sur le site Les Classiques des sciences sociales sont la propriété des **Classiques des sciences sociales**, un organisme à but non lucratif composé exclusivement de bénévoles.

Ils sont disponibles pour une utilisation intellectuelle et personnelle et, en aucun cas, commerciale. Toute utilisation à des fins commerciales des fichiers sur ce site est strictement interdite et toute rediffusion est également strictement interdite.

**L'accès à notre travail est libre et gratuit à tous les utilisateurs. C'est notre mission.**

Jean-Marie Tremblay, sociologue

Fondateur et Président-directeur général,

LES CLASSIQUES DES SCIENCES SOCIALES.

Un document produit en version numérique par Réjeanne Toussaint, bénévole, Chomedey, Ville Laval, Qc. courriel: [rtoussaint@aei.ca](mailto:rtoussaint@aei.ca).

[Page web](http://classiques.uqac.ca/inter/benevoles_equipe/liste_toussaint_rejeanne.html) dans Les Classiques des sciences sociales :

<http://classiques.uqac.ca/inter/benevoles_equipe/liste_toussaint_rejeanne.html>

à partir du texte de :

Michel de Sève

“***La mesure des effets d’inégalité et de discrimination avec des rapports purifiés***.”

In ouvrage sous la direction de André TURMEL, avec la collaboration de Claude Bariteau et Gilles Pronovost, **Chantiers sociologiques et anthropologiques.** Actes du 58e colloque de l’ACSALF 1990, pp 197-230. Montréal : Les Éditions du Méridien, 1993, 274 pp.

La présidente de l’ACSALF, Mme Marguerite Soulière, nous a accordé le 20 août 2018 l’autorisation de diffuser en accès libre à tous ce livre dans Les Classiques des sciences sociales.

Boite_aux_lettres_clair Courriels : La présidente de l’ACSALF, Marguerite Soulière : professeure, École de Service sociale, Université d’Ottawa : [marguerite.souliere@uOttawa.ca](mailto:marguerite.souliere@uOttawa.ca)

André Turmel : [andre.turmel@soc.ulaval.ca](mailto:andre.turmel@soc.ulaval.ca)

Gilles Pronovost : [gilles.pronovost@uqtr.ca](mailto:gilles.pronovost@uqtr.ca)

Police de caractères utilisés :

Pour le texte: Times New Roman, 14 points.

Pour les notes de bas de page : Times New Roman, 12 points.

Édition électronique réalisée avec le traitement de textes Microsoft Word 2008 pour Macintosh.

Mise en page sur papier format : LETTRE US, 8.5’’ x 11’’.

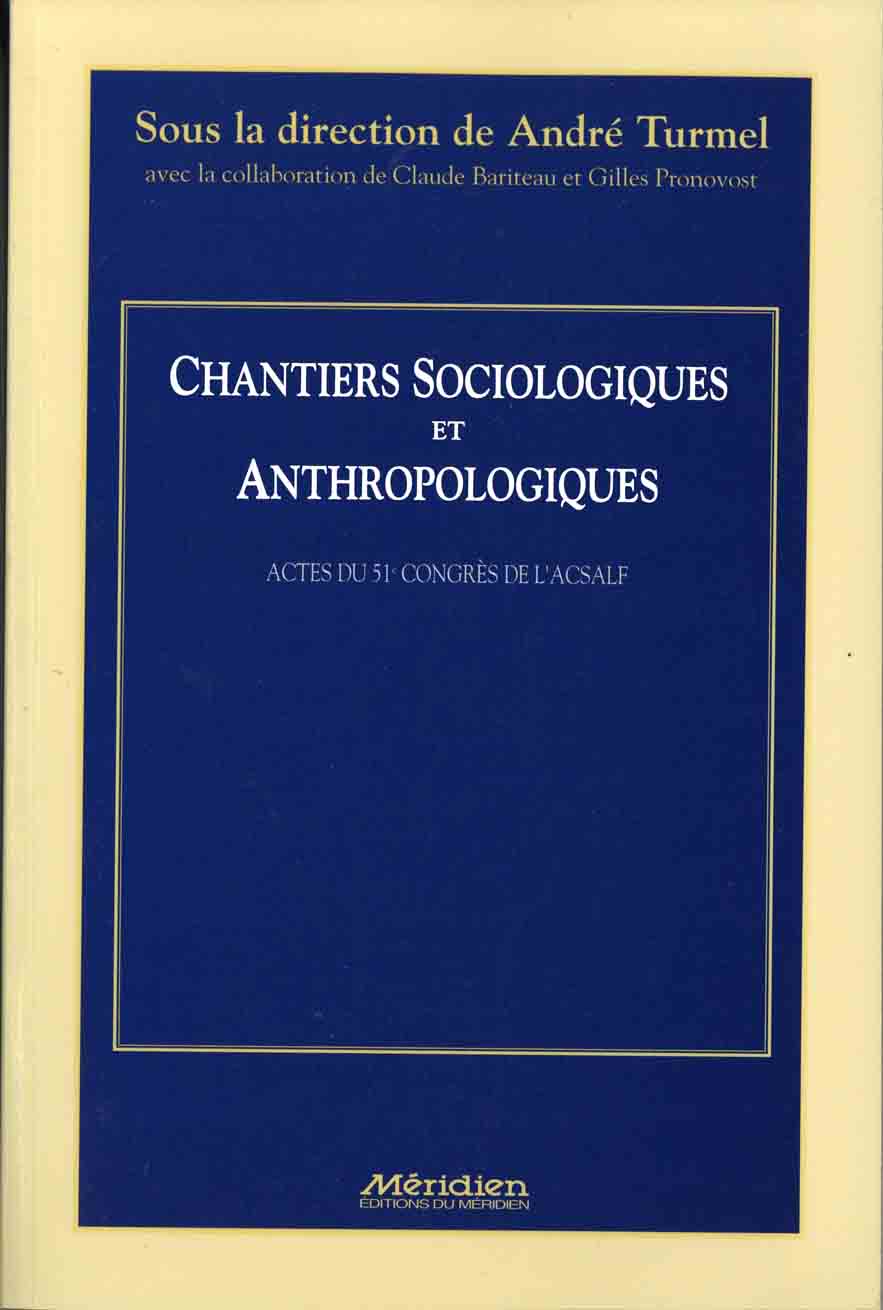
Édition numérique réalisée le 22 juin 2020 à Chicoutimi, Québec.

fait_sur_mac

Michel de Sève

Sociologue, professeur, département de sociologie, Université Laval

La mesure des effets d’inégalité et de discrimination  
avec des rapports purifiés



In ouvrage sous la direction de André TURMEL, avec la collaboration de Claude Bariteau et Gilles Pronovost, **Chantiers sociologiques et anthropologiques.** Actes du 58e colloque de l’ACSALF 1990, pp 197-230. Montréal : Les Éditions du Méridien, 1993, 274 pp.



La présidente de l’ACSALF, Mme Marguerite Soulière, nous a accordé le 20 août 2018 l’autorisation de diffuser en accès libre à tous ce livre dans Les Classiques des sciences sociales.

Boite_aux_lettres_clair Courriel :

La présidente de l’ACSALF, Marguerite Soulière : professeure, École de Service sociale, Université d’Ottawa : [marguerite.souliere@uOttawa.ca](mailto:marguerite.souliere@uOttawa.ca)

**Note pour la version numérique** : La numérotation entre crochets [] correspond à la pagination, en début de page, de l'édition d'origine numérisée. JMT.

Par exemple, [1] correspond au début de la page 1 de l’édition papier numérisée.

[197]

**Chantiers sociologiques et anthropologiques.**

Actes du 58e colloque de l’ACSALF 1990.

“La mesure des effets  
d’inégalités et de discrimination  
avec des rapports purifiés.”

Par Michel de SÈVE

[198]

L'auteur de ce texte tient à remercier M. François Béland pour ses commentaires sur la première version de ce texte et le Service de Consultation Statistique du département de mathématiques et de statistiques de l’Université Laval, en particulier M. Gaétan Daigle de ce service. Évidemment, les erreurs contenues dans ce texte sont le seul fait de l’auteur.

[199]

1. Introduction

En 1986, nous avons proposé un modèle permettant de décrire l’évolution des inégalités et de la discrimination professionnelles selon le sexe et la langue (Béland et de Sève, 1986). Cette communication présente une simplification et une généralisation de ce premier modèle d’analyse.

Le nouveau modèle d’analyse proposé constitue une simplification parce que le premier exigeait la définition de contrastes difficiles à reproduire avec les logiciels existants : le nouveau modèle proposé permet d’obtenir plus facilement les mêmes résultats en utilisant un modèle log-linéaire de type logit. Cette communication généralise également le modèle de départ à des situations où les variables associées aux inégalités ou à la discrimination ne sont plus nécessairement dichotomiques, mais polytomiques. Enfin, s’inspirant des travaux de Clogg et Eliason (1988) et de Santi (1989), la communication propose différents rapports « purifiés » permettant de mieux décrire les inégalités et la discrimination.

2. Les inégalités

Dans le modèle originel (Béland et de Sève, 1986, page 314), l’absence d’inégalités professionnelles entre les sexes étaient définies comme l’absence de différences entre leurs distributions professionnelles une fois éliminée l’influence de la taille des divers sous-groupes comparés, des catégories professionnelles et de l’évolution dans le temps de ces deux facteurs :

[CM (F) – CF (F)] + ICM (A) – CF (A)] = 0

[200]

(dans cette équation, CM et CF désignent les conjoints masculins et féminins de l’exemple présenté alors, et F et A, le groupe linguistique de ceux-ci.)

En d’autres termes, le modèle définit comme des inégalités les différences sur une même variable O entre divers sous-groupes d’une même population. Les différents sous-groupes sont définis selon deux dimensions analytiquement distinctes S et L. En plus, ces différences peuvent être affectées par une variable contextuelle A (dans l’article déjà cité, l’année du mariage).

Le but de ce type d’analyse est de déterminer l’ampleur des différences sur la variable O entre les divers sous-groupes que permet de distinguer une même variable, S par exemple, en éliminant l’effet des différences sur O selon l’autre variable, L dans ce cas. En termes plus concrets, nous cherchons à déterminer les différences professionnelles entre les sexes en éliminant de ces différences celles attribuables aux différences linguistiques. Ces différences professionnelles entre les sexes peuvent évidemment varier d’une année à l’autre.

Finalement, et de la façon la plus générale, il s’agit de déterminer les différences sur une variable dépendante O affectée par deux variables indépendantes S et L : le modèle tente de décrire l’influence « pure » de chacune de ces variables indépendantes (par exemple, S) en éliminant l’influence de l’autre (par exemple, L) sur la variable dépendante et en tenant compte d’une troisième variable en quelque sorte exogène (A).

2.1. Le modèle des inégalités

Si nous considérons toutes les relations et interactions possibles entre les quatre variables O, S, L et A, un modèle log-linéaire saturé contient 16 termes distincts :

TAU, O, S, L, A, OS, OL, OA, SL, SA, LA,

OSL, OSA, OLA, SLA, OSLA

(TAU indique le terme de la moyenne)

Dans un premier temps, ces 16 termes peuvent être regroupés en deux groupes principaux si nous utilisons la logique des modèles logis et considérons la variable O comme la variable dépendante :

[201]

1- 8 termes n’impliquant pas la variable O et mesurant les relations entre les variables S, L et A :

TAU, A, S, SA, L, LA, SL, SLA

2- 8 termes impliquant la variable O que l’on peut répartir en quatre sous-groupes :

- 2 termes permettant de décrire la distribution de la variable dépendante en général et selon les contextes distingués par A :

O et OA

- 2 termes liant les variables O et S selon que l’on tient compte ou non de A :

OS et OSA

- 2 termes liant les variables O et L selon que l’on tient compte ou non de A :

OL et OLA

- 2 termes dits d’interaction liant O, S et L selon que l’on tient compte ou non de A :

OSL et OSLA

Pour détecter et mesurer les inégalités, ce sont les trois dernières paires de termes que nous proposons de retenir :

- le terme OS tient compte des différences sur O entre les sous-groupes distingués par S et le terme OSA, de ces différences selon les contextes ; ces différences sont indépendantes des effets de L sur O : si le terme OS est « nul », il n’y a pas d’inégalités entre les sous-groupes distingués par S ;

- le terme OL tient compte des différences sur O entre les sous-groupes distingués par L et le terme OLA, de ces différences selon les contextes ; ces différences sont indépendantes des effets de S sur O : si le terme OL est « nul », il n’y a pas d’inégalités entre les sous-groupes distingués par L ;

- le terme OSL tient compte des différence sur O entre les sous-groupes distingués par S selon les valeurs de L (ou vice-versa : les différences sur O entre les sous-groupes distingués par L selon les valeurs de S) ; le terme OSLA indique si ces différences varient selon les contextes définis à l’aide [202] de A ; en d’autres termes, ces termes mesurent les effets d’interaction des inégalités selon S et L ; s’ils sont nuis, les inégalités selon S sont les mêmes d’un sous-groupe de L à l’autre.

Ce modèle logis, particulièrement les paires de termes (OS, OSA), (OL, OLA), (OSL, OSLA) permet de mener les mêmes analyses que les contrastes du modèle original proposé en 1986. Les chis deux observés sont les mêmes et les différents coefficients d’inégalités obtenus sont égaux à une constante multiplicative près.

2.2. Première étape :  
trouver un modèle satisfaisant

Comme pour toute analyse log-linéaire, il s’agit de déterminer les termes nécessaires pour rendre compte des fréquences observées avec suffisamment de précision (chi deux du modèle). Dans les différents modèles, et en s’inspirant de la logique des modèles logis, on distinguera deux types d’effets regroupant les 16 termes possibles du modèle :

1- ce que l’on pourrait appeler des effets structuraux mesurant d’une part les relations entre les variables S, L et A et, d’autre part, la variable O et ses variations selon les contextes :

TAU, A, S, SA, L, LA, SL, SLA, O et OA

2- des effets d’inégalités : des effets dits « simples » :

OS, OSA, OL, OLA

et des effets dits « d’interaction » :

OSL et OSLA

De façon générale, nous introduirons dans tous les modèles testés les 10 effets dits structuraux et nous tenterons d’éliminer autant d’effets d’inégalités que possible. Comme le modèle contient 6 effets d’inégalités, il existe une possibilité de 64 modèles (non hiérarchiques) et il y a avantage à développer une stratégie de recherche d’un modèle à la fois satisfaisant et le plus simple possible.

En premier lieu, on aura avantage à éliminer particulièrement les effets d’interaction OSL et OSLA pour deux raisons :

1- la nécessité de ces termes signifie que les effets d’inégalités d’une variable varient selon les valeurs de l’autre variable [203] pouvant affecter les inégalités ; dans une telle situation, présenter les inégalités selon la première variable sans tenir compte de la seconde peut représenter une simplification abusive de la réalité (Santi, 1989, pages 385 et 386),

2- ceux-ci compliquent le calcul des effets d’inégalités si les rapports d’inégalités dits « additifs » sont utilisés (voir la section 2.4).

Tenter d’éliminer en premier ces termes d’interaction est évidemment un choix arbitraire car, statistiquement, les termes OSA et OLA sont aussi complexes que le terme OSL : si on désire obtenir une description des inégalités indépendante des contextes définis par A, on pourra essayer d’éliminer dans un premier temps les termes OSA et OLA). On testera dont les trois modèles suivants (sans oublier d’introduire les dix termes des effets structuraux) :

1) OS, OSA, OL, OLA, OSL

2) OS, OSA, OL, OLA

3) (sans aucun effet d’inégalités, mais avec les effets structuraux).

Le modèle # 1 est le modèle non saturé le plus simple ne contenant pas l’effet d’inégalités d’interaction le plus complexe : OSLA. Le modèle #2 est le modèle le plus complexe ne contenant que des effets simples d’inégalités. Enfin, le modèle #3 ne contient aucun effet d’inégalités, simples ou d’interaction : la valeur de son chi deux indique l’ampleur des variations expliquées par les effets d’inégalités.

Les termes OSL et OSLA peuvent être éliminés pour des raisons statistiques ou pratiques.

*Statistiquement*, le terme OSLA ne sera pas nécessaire si le chi deux du modèle #1 est inférieur au seuil de signification choisi ; le terme OSL pourra être éliminé si le chi deux du modèle #2 est non significatif statistiquement.

*Pratiquement*, les termes OSLA et/ou OSL pourront être éliminés du modèle même s’ils sont statistiquement nécessaires si le modèle 2 explique la presque totalité des variations attribuables aux effets d’inégalités. Pour établir la proportion des variations attribuables aux inégalités expliquées par les effets simples d’inégalités (OS, OSA, OL et OLA) avec ou sans (OSL), on comparera les variations totales [204] expliquées par les effets d’inégalités aux variations expliquées par les effets simples avec ou sans (OSL) :

|  |  |
| --- | --- |
| R2 = | chi deux (modèle #3) — chi deux (modèle #1 ou #2) |
| chi deux (modèle #3) |

(D’autres mesures d’ajustement sont décrites par Hagenaars, 1990, pages 64 à 68.)

Si les termes OSL et/ou OSLA peuvent être éliminés, on cherchera à établir si les modèles suivants suffisent pour expliquer les données :

4) OS, OSA, OL,

5) OS, OL, OLA

6) OS, OL

7) OS

8) OL

En comparant les modèles #4, #5 et #6 au modèle #2, on cherchera à déterminer si les termes OSA et/ou OLA sont nécessaires. De façon similaire, les modèles #7 et #8 comparés au modèle #6 permettront de déterminer si les effets d’inégalités OS et OL sont nécessaires.

Si le terme OSL ne peut être éliminé, on cherchera parmi les modèles suivants afin d’éliminer les effets contextuels des effets simples :

9) OS, OSA, OL, OSL

10) OS, OL, OLA, OSL

11) OS, OL, OSL

12) OL, OLA, OSL (modèle non hiérarchique)

13) OS, OSA, OSL (modèle non hiérarchique)

Les modèles #9, #10 et #11 permettront de déterminer si l’effet de la variable contextuelle A est nécessaire parmi les effets simples et les modèles #12 et #13 donneront une idée de l’importance des effets simples en les comparant au modèle #1.

En même temps, on aura sans doute avantage à décomposer les variations totales des fréquences observées selon qu’elles reflètent des effets structuraux ou des effets d’inégalités et, dans les effets d’inégalités, les variations attribuables aux effets d’interaction (OSL et OSLA) des variations attribuables aux effets simples (OS, OSA, OL et OLA).

[205]

Le fait que cette stratégie mette l’accent sur les effets dits d’inégalités ne signifie pas que les effets structuraux sont sans intérêt. On aura donc avantage à examiner l’importance de ceux-ci, notamment :

- O et OA pour décrire les variations de O et son évolution ou ses différences selon les contextes,

- S, L, SL (et SA, LA, SLA) pour décrire l’importance des divers sous-groupes distingués dans la population (et l’évolution de celle-ci).

2.3. Deuxième étape : décrire les inégalités

2.3.1. Comment décrire les inégalités ?

Pour décrire les inégalités, nous proposons de comparer les fréquences théoriques des répondants occupant une même occupation une même année d’un groupe à celle d’un autre groupe dit de référence. Par exemple, pour l’année c (A=c) et chaque occupation a (0=a), le rapport suivant sera utilisé pour décrire les inégalités selon le sexe :

|  |  |  |
| --- | --- | --- |
| F (OA, ac) | = | Femmes (OA, ac) |
| H (OA, ac) | Hommes (OA, ac) |

Si ce rapport est supérieur à l’unité, les femmes sont plus nombreuses que les hommes dans l’occupation a pour l’année c. Par contre, s’il est inférieur à l’unité, les hommes sont plus nombreux que les femmes dans la même occupation. Enfin, un rapport égal à l’unité indique une égalité entre les sexes dans cette occupation.

Cependant, même s’il y a égalité entre les sexes (ou les groupes ethniques), on ne peut s’attendre à obtenir des coefficients égaux à 1 car les fréquences théoriques brutes sont influencées par d’autres facteurs que les inégalités comme le nombre de femmes ou d’hommes présents dans l’échantillon chaque année. Avant d’examiner ces rapports, nous « purifierons » donc ces rapports en nous inspirant des travaux de Clogg et Eliason, et de Santi (Clogg et Eliason, 1988 ; Santi, 1989).

De plus, les fréquences d’hommes et de femmes dans une même occupation pour une même année sont en fait composé des fréquences [206] d’hommes et de femmes dans la même occupation pour la même année dans chaque sous-groupe défini par l’autre variable pouvant affecter les inégalités. Ainsi, il y a des hommes francophones, anglophones et allophones comme il y a des femmes francophones, anglophones et allophones. La question est donc comment intégrer ces différents sous-groupes linguistiques au rapport femmes/hommes de telle façon que sa valeur indique l’existence, l’ampleur et la direction des inégalités.

Nous avons examiné trois possibilités de rapports. Pour continuer avec le cas des inégalités sexuelles :

a) *des rapports additifs*:

en sommant les fréquences théoriques (purifiées) des différents groupes linguistiques (par exemple, les francophones et anglophones si seulement deux groupes linguistiques sont comparés) de chaque sexe :

|  |  |  |
| --- | --- | --- |
| F (OA, ac) | = | F (OLC, alc) + F (OLC, a2c) |
| H (OA, ac) | H (OLC, alc) + H (OLC, a2c) |

b) *des rapports de rapports*:

en divisant le rapport femmes/hommes chez les francophones (L=l) par le même rapport chez les anglophones (L=2) :

|  |  |  |
| --- | --- | --- |
| F (OA.ac) | = | (F (OLC, alc) / H(OLC, alc)) |
| H (OA, ac) | (F (OLC, a2c) / H (OLC, a2c)) |

c) *des rapports multipliés*:

en multipliant la série des rapports femmes/hommes des différents groupes linguistiques :

|  |  |  |
| --- | --- | --- |
| F (OA, ac) | = | F (OLC, alc) F (OLC, a2c) |
| H (OA, ac) | H (OLC. alc) H (OLC, a2c) |

L’examen du deuxième type de rapports révèle que ceux-ci ne dépendant que des termes OSL et OSLA (à la puissance 4 dans le cas de variables S et L dichotomiques) et aucunement des termes OS et OSA : ils indiquent donc une absence d’inégalités quand il n’y a pas d’interaction entres les inégalités et sont donc à rejeter.

[207]

Par contre, les rapports du premier type et du troisième dépendent des termes OS et OSA ; ils sont égaux à l’unité quand il y a absence d’inégalités (et d’interaction). Comme permet de le constater l’annexe, les rapports du troisième type sont même plus facile à calculer, en particulier quand S et/ou L ne sont pas dichotomiques. Nous avons cependant choisi de privilégier les rapports du premier type car leur définition nous semble plus proche de la notion habituelle de rapport entre deux quantités.

2.3.2. Les rapports additifs

Dans cette section, nous montrerons comment les rapports additifs peuvent être obtenus en utilisant les termes du modèle retenu à la première étape de cette section. Pour simplifier les démonstrations, nous utiliserons la formalisation multiplicative des modèles log-li- néaires, quitte à effectuer les transformations logarithmiques nécessaires à la fin. Nous présenterons donc les formules nécessaires pour estimer les rapports du type :

|  |  |
| --- | --- |
|  | Femmes (ac) / Hommes (ac) |
| et |  |
|  | Francophones (ac) / Anglophones (ac) |

Le fait que les variables O et A soient polytomiques n’affecte évidemment pas ces résultats. Par contre, il est apparu que le fait que les variables S et L le soient a des conséquences sur le calcul des coefficients désirés. Nous distinguerons donc quatre situations selon que ces deux variables sont dichotomiques ou non.

Dans ces formules, nous avons conservé, lorsqu’il était nécessaire, les termes OSL et OSLA même si dans une telle situation, il peut être douteux de présenter les inégalités selon une variable sans tenir compte de l’autre.

Notation :

|  |  |
| --- | --- |
| OSLA (osla) = | le terme multiplicatif d’un modèle pour la catégorie o de la variable O, la catégorie s de la variable S, la catégorie 1 de la variable L et la catégorie a de la variable A |
| F (osla) = | la fréquence théorique de la cellule de la catégorie o de la variable O, la catégorie s de la |

[208]

|  |  |
| --- | --- |
|  | variable S, la catégorie 1 de la variable L et la catégorie a de la variable A |
| F (osla)\* = | la même fréquence théorique, mais « purgée » de certains effets (voir plus loin) |

Dans le cas des inégalités sexuelles, on désire obtenir le rapport femmes/hommes dans une même occupation « a » et pour une même année « c ». Ces fréquences théoriques réunissent en fait, pour chaque sexe, les francophones (L=l) et les anglophones (L=2) d’un même sexe :

|  |  |  |
| --- | --- | --- |
| Femmes (ac) | = | F (al le) + F (a 12c) |
| Hommes (ac) | F (a21c) + F (a22c) |

De façon similaire, pour les inégalités linguistiques, on peut désirer obtenir le rapport francophones/anglophones pour une même occupation « a » et une même année « c ». Ces fréquences théoriques réunissent en fait, pour chaque groupe linguistique, les femmes (S=l) et les hommes (S=2) d’un même groupe linguistique :

|  |  |  |
| --- | --- | --- |
| Francophones (ac) | = | F (al le) + F (a21c) |
| Anglophones (ac) | F (al2c) + F (a22c) |

Telles quelles, les fréquences théoriques de ces rapports reflètent non seulement les effets d’inégalités, mais également les effets structuraux :

|  |  |
| --- | --- |
| F (ai je) = | tau\*O (a) \*A(c) \*OA (ac) \* OS (ai) \*OSA (aie) \*OL (aj) \*OLA(ajc) \* OSL (aij) \*OSLA (aijc) \* S (i) \*SA (ia) \*L (j) \*SA (je) \*SL (ij) \*SLA (ijc) |

Avant de calculer les rapports, ces fréquences théoriques seront purifiées de deux types d’effets :

- des effets structuraux : S, SA, L, LA, SL, SLA, O, A, OA ;

- pour obtenir les coefficients des effets d’inégalités sexuelles, des effets d’inégalités linguistiques OL et OLA ; par contre, pour obtenir les coefficients des effets d’inégalités linguistiques, des effets d’inégalités sexuelles OS et OSA.

Par « purifier », nous entendons « diviser les fréquences théoriques par les termes correspondant aux effets à éliminer » avant de les réunir [209] et de les mettre en rapport. Par exemple, pour obtenir les inégalités sexuelles :

|  |  |
| --- | --- |
| F (aijc) \* = | tau\*O (a) \*A (c) \*OA (ac) \*  OS (ai) \*OSA (aie) \*OL (aj) \*OLA (ajc) \*  OSL (aij) \*OSLA (aijc) \*  S (i) \*SA (ia) \*L (j) \*SA (je) \* SL (ij) \*SLA \*(ijc) |
|  | (tau\*O (a) \*A(c) \*OA(ac) \*  OL (aj) \*OLA (ajc) \*  S(i) \*SA(ia) \*L(j) \*LA(jc) \*SL(ij) \*SLA(ijc) |

Cette purification permet d’obtenir deux grands types de rapports :

1- le rapport de la fréquence théorique dans une catégorie de S sur la fréquence théorique d’être dans une autre catégorie de S ; si S est polytomique, on comparera les différentes catégories de S à une même catégorie de S ; dans les deux cas, les fréquences sont purifiées des effets suivants :

tau

OL et OLA

O, A et OA

S et SA

L et LA

SL et SLA

ce rapport de deux fréquences théoriques purifiées sera désigné par l’acronyme RIS :

|  |  |  |  |
| --- | --- | --- | --- |
| RIS = | OS (al) \*OSA (alc) | \* | Somme (j) (OSL (alj) \*OSLA (aljc) |
| OS (a2) \*OSA(a2c) | Somme (j)(OSL (a2J) \*OSLA (a2jc) |

2- le rapport de la fréquence théorique dans une catégorie de L sur la fréquence théorique d’être dans l’autre catégorie de L ; si L est polytomique, on comparera les différentes catégories de L à une même catégorie de L ; dans les deux cas, les fréquences sont purifiées des effets suivants :

tau

OS et OSA

O, A et OA

S et SA

L et LA

[210]

SL et SLA

ce rapport de deux fréquences théoriques purifiées sera désigné par l’acronyme RIL :

|  |  |  |  |
| --- | --- | --- | --- |
| RIL = | OL (al) \*OLA (alc) | \* | Somme (i)(OSL(ail) \*OSLA (ailc) |
| OL (a2) \*OLA(a2c) | Somme (i) (OSL(ai2) \*OSLA(ai2c) |

Note :

Dans les formules qui suivent, on supposera un modèle saturé incluant tous les termes possibles. Si le modèle retenu n’est pas saturé, on remplacera les termes non retenus par la valeur 1.

Cas 1 : *S et L sont dichotomiques*:

Il n’y a pas de différence selon qu’il y a ou non interaction entre les inégalités (OSL et OSLA) ; ces termes disparaissent en effet algébriquement des rapports recherchés (la catégorie 1 de la variable S désigne les femmes) :

RIS = (OS (al))2 \* (OSA (alc))2

RIL = (OL (al))2 \* (OLA(alc))2

Si nous transformons logarithmiquement de ces deux rapports, nous obtiendrons (à la constante 2 près) les coefficients « logits » de l’article (Béland et de Sève, 1986, pages 325 à 324). Avec des résultats log-linéaires sous forme logarithmique, il faudra donc sommes les termes du modèle :

coefficient inég. sex. = 2 \* (OS (al) + OSA (alc)

coefficient inég. lin. = 2 \* (OL (al) + OLA (alc)

On reconnaîtra là tout simplement les coefficients logits d’un modèle log-linéaire employé pour faire ce type d’analyse. Quand l’une des deux variables n’est pas dichotomique, cette forme des coefficients est moins claire comme le montreront les sections suivantes.

Cas 2 : *S est dichotomique et L est polytomique*:

Il y a une différence selon que l’on suppose ou non un effet d’interaction (OSL et OSLA) ; de plus, le rapport RIL ne peut être simplifié.

Pour la variable L, la catégorie 2 désigne une catégorie de référence choisie par l’analyste (par exemple, le groupe ethnique le [211] plus favorisé ou majoritaire); par contre, la catégorie 1 de L désigne les différents sous-groupes selon L que l'on veut comparer au groupe le plus favorisé ou majoritaire: de sous-groupe en sous-groupe, la catégorie 1 varie donc alors que le sous-groupe 2 demeure toujours le même:

|  |  |
| --- | --- |
| RIS = (OS (a1))2 \* (OSA(a1c))2 \* | Som (i) (OSL (a1j) OSLA (a1jc)) |
| Som (j) (OSL (a2j) OSLA (aa2jc)) |

ou

|  |  |
| --- | --- |
| RIS = (OS (a1)2 \* (OSA (a1c))2 \* | Som (j) (OSL (a1j) OSLA (a1cj)) |
| Som (j) (l/(OSL(a1j)OSLA(a1jc))) |

|  |  |  |  |
| --- | --- | --- | --- |
| RIL = | OL (al) \* OLA (a1c) | \* | Som (i) (OSL (ai1) OSLA(ai1c)) |
| OL (a2) \* OLA (a2c) | Som (i) (OSL (ai2) OSLA(ai2c)) |

Si les termes d'interaction OSL et OSLA sont non significatifs ou exclus pour des raisons pratiques, la deuxième partie des coefficients est égale (ou supposée égale) à 1 et disparaît.

Cas 3: *S est polytomique et L est dichotomique*:

Comme lorsque L est polytomique alors que S ne l'est pas, il y a une différence selon que l'on suppose ou non un effet d'interaction (OSL et OSLA); de plus, le rapport RIS ne peut être simplifié.

Également, pour la variable S, il faudra choisir un sous-groupe de référence (S=2 dans les formules) et les termes propres aux différents autres sous-groupes selon S remplaceront les termes pour lesquels S=l dans ces formules.

|  |  |  |  |
| --- | --- | --- | --- |
| RIS = | OS (a1) \* OSA (a1c) | \* | Som (j) (OSL (a1j) OSLA (a1jc)) |
| OS (a2) \* OSA (a2c) | Som (j) (OSL (a2j) OSLA (a2jc)) |

|  |  |
| --- | --- |
| RIL = (OL (a 1 ))2 \* (OLA (a 1 c))2 | Som (i) (OSL (ai1) OSLA (ailc)) |
| Som (i) (OSL (ai2) OSLA (ai2c)) |

ou

|  |  |
| --- | --- |
| RIL = (OL (al))2\*(OLA (alc))2\* | Som (i) (OSL (ai1) OSLA (ai1c)) |
| Som (i) (l/ (OSL (ai1) OSLA(ai1c)) |

[212]

Cas 4: *S et L sont polytomiques toutes les deux*:

Pour les deux variables S et L, la valeur 2 désigne le sous-groupe de référence et la valeur 1, les différents autres sous-groupes comparés à chacun des sous-groupes de référence:

|  |  |  |  |
| --- | --- | --- | --- |
| RIS = | OS (a1) \* OSA (a1c) | \* | Som (j) (OSL (a1j) OSLA (a1jc)) |
| OS (a2) \* OSA (a2c) | Som (j) (OSL (a2j) OSLA (a2jc)) |

|  |  |  |  |
| --- | --- | --- | --- |
| RIL = | OL (a1) \* OLA (a1c) | \* | Som (i) (OSL (ai1) OSLA (ai1c)) |
| OL (a2) \* OLA (a2c) | Som (i) (OSL (ai2) OSLA (ai2c)) |

3. La mesure de la discrimination

Comme pour les inégalités, nous désirons comparer différents groupes définis par des variables S et L selon une même variable O dans différents contextes indiqués par une variable A. Cependant, nous ne cherchons pas à comparer directement la présence dans chaque O des différents groupes, mais cette présence étant donnée la «présence» de ces groupes dans des occupations «antérieures» ou dans des occupations servant de point de comparaison (variable G, par exemple, dans l'article originel (Béland et de Sève, 1986, les occupations des pères). En d'autres termes, nous voulons comparer, par exemple pour le sexe :

|  |  |  |
| --- | --- | --- |
|  | maintenant | auparavant |
|  | G=l | G=2 |
|  | (conjoints) | (pères des conjoints) |
|  | S=1 S=2 | S=l S=2 |
| 0=1 | diffl | diff2 |
| 0=2 | diff3 | diff4 |

Il n'y a pas de discrimination si les rapports suivants sont égaux à 1 (ou leurs logarithmes égaux à 0):

(diffl/diff2) = 1

(diff3/diff4) = 1

etc.

Dans le contexte de l'article originel (Béland et de Sève, 1986, page 313), ces rapports de rapports sont égaux à zéro (absence de discrimination) si les différences entres les occupations des conjoints au moment du mariage et celles de leurs pères lors de la naissance des conjoints sont égales chez les deux sexes. Ainsi, l'absence de [213] discrimination est définie par l’égalité suivante comparant les occupations des pères et des conjoints en tenant compte de la présence de deux groupes linguistiques :

{[CF (F) – PF (F)] + [CF (A) – PF (A)]} =

{[CM (F) – PM (F)] + [CM (A) – PM (A)]}

(ou CF et CM indiquent les conjoints féminins et masculins, PF et PM, les pères des conjoints féminins et masculins, et, enfin, les lettres F et A entre parenthèses, la langue des conjoints ou des pères.)

3.1. Le modèle de la discrimination

Le modèle doit tenir compte de 5 variables (O, S, L, G et A) et, quand il est saturé, il contient 32 termes. Ceux-ci peuvent se répartir en deux groupes principaux selon l’absence ou la présence de la variable O :

1- 16 effets dits « structuraux » excluant la variable O :

ces termes décrivent les relations entre les variables S, L, G et A :

TAU, S, L, G, A,

SL, SG, SA,

LG, LA, GA,

SLG, SLA, SGA, LGA,

SLGA

certains sont sans doute intéressants en eux-mêmes ; par exemple, les termes SGA et LGA permettent de décrire l’évolution de la présence relative des sexes et des groupes linguistiques selon les années si nous considérons les valeurs de G dénotant les générations étudiées ;

2- 16 effets incluant la variable O :

ces 16 effets se divisent en deux sous-groupes selon qu’ils permettent ou non d’examiner les différences « générationnelles » (G) entre les sexes ou les groupes linguistiques :

a) 10 effets « structuraux » décrivant la distribution de O selon les quatre autres variables S, L, A et G, mais sans distinguer les différentes « générations » ou états de G :

O, OS, OL, OG, OA,

OSL, OSA, OLA, OGA,

OSLA,

[214]

b) 6 effets permettant de décrire la discrimination parce qu’ils mettent en relation les différences occupationnelles selon la « génération » G :

b.1) des effets simples de discrimination entre les groupes distingués par S :

OSG et OSGA

le terme OSG est significatif si les différences entre les S sur O ne sont pas les mêmes selon que G=1 et G=2 (si G est polytomique, voir plus loin) ; si ces différences varient selon le « contexte » A, OSGA est aussi significatif ;

b.2) des effets simples de discrimination entre les groupes distingués par L :

OLG et OLGA

le terme OLG est significatif si les différences entre les L sur O ne sont pas les mêmes selon que G=1 et G=2 ; si ces différences varient selon le « contexte » A, OLGA est aussi significatif ;

b.3) deux effets d’interaction entre les discriminations selon S et L :

OSLG et OSLGA

ces effets sont non « nuis » si les différences entre les S sur O compte tenues des différences sur G varient selon les valeurs de L (ou vice versa) ; de plus, ces différences peuvent varier selon le contexte A.

3.2. Première étape :  
trouver un modèle satisfaisant

Cette étape de l’analyse se fera comme pour les inégalités. Le nombre de termes permettant de contrôler les divers effets « structuraux » est cependant plus grand. Ainsi, nous testerons des modèles

— incluant toujours les 26 effets structuraux : les 16 ne contenant pas la variable O et les 10 la contenant :

TAU, S, L, G, A,

SL, SG, SA,

LG, LA, GA,

SLG, SLA, SGA, LGA,

SLGA

[215]

et

O, OS, OL, OG, OA,

OSL, OSA, OLA, OGA,

OSLA,

— mais faisant varier les 6 termes décrivant la discrimination de façon à obtenir un modèle satisfaisant le plus simple possible :

OSG, OSGA,

OLG, OLGA,

OSLG et OSLGA.

Dans un premier temps, on aura avantage à tenter d’éliminer les termes d’interaction entre les discriminations (OSLG et OSLGA) pour des raisons statistiques (test du chi deux) ou pour des raisons pratiques (peu d’importance de ces termes pour expliquer les variations des données). Les modèles suivants permettront d’évaluer la nécessité des termes OSLG et OSLGA :

1) OSG, OSGA, OLG, OLGA, OSLG

2) OSG, OSGA, OLG, OLGA

3) (sans effets de discrimination, mais incluant les 26 effets structuraux comme les modèles 1 et 2)

Le degré de signification du modèle 1 permettra de déterminer si le terme absent de ce modèle OSLGA est réellement nécessaire et, son chi deux comparé à celui du modèle 3, s’il est pratiquement important pour bien reproduire les fréquences observées.

Par contre, le degré de signification du modèle 2 (ou la différence entre les chis deux des modèles 1 et 2) permettra d’évaluer la nécessité du terme OSLG et son chi deux comparé à celui du modèle 3, l’importance pratique de celui-ci.

Si les termes OSLG et OSLGA peuvent être éliminés, on cherchera à établir l’utilité des effets « simples » de discrimination en examinant les modèles suivants (ou une suite similaire) :

4) OSG, OSGA, OLG

5) OSG, OLG, OLGA

6) OSG, OLG

7) OSG

8) OLG

[216]

La comparaison des modèles 4 et 5 au modèle 2 permettra de déterminer si l’on doit tenir compte du contexte A (les années dans l’article déjà cité) ou non pour rendre compte des données. Si la réponse est négative pour les deux types de discrimination (selon S et selon L), le modèle 6 devrait être non significatif. Enfin, les modèles 7 et 8 permettront de déterminer s’il est nécessaire de tenir compte des deux types de discrimination, d’une seule ou d’aucune.

S’il apparaît possible d’obtenir un modèle satisfaisant sans aucun des effets de discrimination (OSG, OLG, OSGA, etc.), les différences sur O selon S et L sont invariantes selon les différents états de G : a priori, on cherchera sans doute leur source dans un premier temps parmi les 10 effets structuraux impliquant O décrits dans la section 3.1.

Par contre, s’il est impossible d’éliminer le terme OSLG ou le terme OSLGA, il faudra recourir à des modèles non hiérarchiques comme ceux présentés à la fin de la section 2.2 pour essayer de simplifier autant que possible le modèle de discrimination. Comme pour les inégalités, la présentation des rapports de discrimination (voir la section 3.4) selon un seul des facteurs de discrimination risque alors de donner une image trop simplifiée de la réalité.

3.3. Seconde étape :  
décrire la discrimination

Comme pour les inégalités, il s’agit de définir un rapport permettant de comparer les fréquences de deux groupes ; cependant ce rapport en est un plus complexe car il ne prétend pas uniquement refléter la présence relative des deux groupes dans, par exemple, une occupation, mais cette présence relative selon une situation antérieure ou une situation servant de point de comparaison.

Ainsi, dans l’article, pour la discrimination sexuelle, la présence des femmes et des hommes dans chaque occupation est comparée à la présence de leurs pères dans la même occupation :

|  |
| --- |
| Conjoints féminins / Conjoints masculins |
| Pères des conjoints fém. / Pères des conjoints masc. |

[217]

Ce rapport est égal à l’unité si la présence relative des conjoints est égale à celle de leurs pères respectifs. Dans un tel cas, on peut prétendre que les différences entre conjoints sont semblables à celles de leurs origines et qu’il n’y a pas de discrimination.

(Ce rapport est équivalent au rapport présenté au début de cette partie sur la discrimination :

|  |
| --- |
| Conjoints féminins / Pères des conjoints féminins |
| Conjoints masculins / Pères des conjoints masculins) |

Si, comme dans l’article, G désigne la génération (la valeur 1 pour les conjoints et la valeur 2 pour leurs pères), le rapport à purifier pour la discrimination sexuelle aura la forme suivante pour une occupation a et une « année » c :

|  |
| --- |
| F (OSGA, a11c)/F (OSGA, a21c) |
| F (OSGA, a12c)/F (OSGA, a22c) |

(ce qui peut se lire aussi :

|  |
| --- |
| F (OSGA, a11c)/F (OSGA, a12c) |
| F (OSGA, a21c)/F (OSGA, a22c) ) |

Comme pour les inégalités, ces fréquences réunissent les répondants appartenant aux différents groupes que permettent de distinguer les valeurs de la variable L : pour les femmes (S=l) comme pour les hommes (S=2), elles s’écriront pour les différentes valeurs k de G :

|  |
| --- |
| F (OSGA, a1kc) – Som (j) F (OSLGA, a1jkc) |
| F (OSGA, a2kc) – Som (j) F (OSLGA, a2jkc) |

Avant de définir le rapport de discrimination sexuelle, chacune des fréquences sommées sont « purifiées » des termes du modèle n’impliquant pas des effets de discrimination sexuelle simple ou d’interaction de telle sorte qu’elles s’écrivent en général de la façon suivante pour une occupation a donnée et une année c particulière :

F90SGA, aikc)\* =

Som (j) OSG (aik) \*OSGA (aikc) \*OSLG (aijk) \*OSLGA (aijkc)

ou, un peu mieux,

[218]

F (OSGA, aikc) \* =

OSG (aik) \*OSGA (aikc) \* Som (j) (OSLG (aijk) \*OLGA (aijkc))

De la même façon, et toujours dans le contexte de l’article, pour comparer les groupes linguistiques, nous comparerons les occupations des francophones à celles des anglophones en tenant compte des occupations de leurs pères :

|  |
| --- |
| Conjoints francophones / Conjoints anglophones |
| Pères des conjoints franc. / Pères des conjoints angl. |

ou :

|  |
| --- |
| Conjoints francophones / Pères des conjoints franc. |
| Conjoints anglophones / Pères des conjoints angl. |

Les fréquences des francophones et des anglophones réunissent les conjoints ou les pères des conjoints des deux sexes : pour les francophones, elles ont la forme :

F (OLGA, a1kc) = Som (i) F (OSLGA, ai1kc)

et, pour les anglophones :

F (OLGA, a2kc) = Som (i) F (OSLGA, ai2kc)

Avant d’être utilisées pour définir le rapport de discrimination linguistique, ces fréquences sont cependant purifiées des termes n’impliquant pas les effets de discrimination linguistique simples ou d’interaction ; toujours pour la même occupation a et la même année c, elles ont la forme suivante :

F (OLGA, alkc)\* =

Som (i) OLG(a1k)\*OLGA(a1kc)\*OSLG(ai1k)\*OLGA(ai1kc)

ou, un peu mieux,

F (OLGA, alkc)\* =

OLG (a1k)\*OSGA (a1kc) \* Som(i) (OSLG(ai1k)\*OLGA(aikc))

Dans les situations suivantes, RDS indique le rapport de la discrimination sexuelle alors que RDL indique le rapport de la discrimination linguistique. Les formules sont présentées en supposant que le modèle log-linéaire est calculé sous sa forme multiplicative. Il est relativement facile de faire les transformations nécessaires pour un modèle sous forme logarithmique.

[219]

Comme pour les inégalités, il faut distinguer quatre situations selon que S et/ou L sont dichotomiques ou non. De plus, et de façon similaire, l’absence des termes d’interaction OSLG et OSLGA dans le modèle retenu simplifie les coefficients recherchés.

La valeur 1 de la variable S correspond au sexe féminin dans le contexte de l’article et la valeur 2, au sexe masculin ; si S est polyto-mique, 1 peut varier selon les groupes comparés, mais 2, demeure fixe et désigne le groupe de référence pour la variable S. De la même façon, la valeur 1 de la variable L désigne les francophones ou un groupe à comparer et la valeur 2, les anglophones ou le groupe de référence pour la variable L.

De façon similaire, la valeur 1 de la variable G indique l’occupation ou la génération comparée (les occupations des conjoints dans l’article) et la valeur 2 de la même variable, l’occupation ou la génération servant de point de référence pour établir la discrimination (les occupations des pères dans l’article). Les rapports RDS et RDL étant plus simples quand G est dichotomique, dans chaque cas, ils sont présentés selon que G est dichotomique ou non.

Comme pour les inégalités, il est possible de définir les rapports RDS et RDL par le produit des rapports propres à chaque sous-groupe. Par exemple, RDS pourrait être défini de la façon suivante :

|  |  |  |
| --- | --- | --- |
| Conjoints fém. francophones | \* | Conjoints fém. anglophones |
| Conjoints mas. francophones | Conjoints mas. anglophones |
| Pères des conj. fém. franc | Pères des conj. fém. angl. |
| Pères des conj. mas, franc. | Pères des conj. mas. angl. |

et le rapport RDL également par le produit des rapports selon le sexe:

|  |  |  |
| --- | --- | --- |
| Conjoints fém. francophones | \* | Conjoints mas. francophones |
| Conjoints fém. anglophones | Conjoints mas. anglophones |
| Pères des conj. fém. franc. | Pères des conj. mas. franc. |
| Pères des conj. fém. angl. | Pères des conj. mas. angl. |

[220]

Les rapports RDS et RDL définis ainsi présentent l’avantage d’être beaucoup plus simples que ceux définis additivement car les termes OSLG et OSLGA se simplifient dans tous les cas. L’annexe contient les formules de ces rapports multiplicatifs selon les différents cas distingués. Cette possibilité de simplification algébrique ne signifie cependant pas qu’il soit plus justifié d’employer ces rapports si les termes OSLG et/ou OSLGA sont nécessaires pour expliquer les fréquences observées.

Cas 1 : *S et L sont dichotomiques :*

Les termes d’interaction disparaissent des rapports grâce à des simplifications algébriques.

a) G est dichotomique

RDS = OSG (al l)4 \* OSGA (al le)4

RDL = OLG (all)4 \* OLGA (allc)4

À la constante « 4 » près, ces deux expressions permettent de retrouver les coefficients de l’article après transformation logarithmique.

b) G est polytomique

RDS = (OSG (al 1) \* OSGA (al le) \* OSG (a22) \* OSGA (a22c)2

RDL = (OLG (al 1) \* OLGA (al le) \* OLG (a22) \* OLGA (a22c)2

Cas 2 : *S est dichomotime et L est polytomique*:

a) G est dichotomique

RDS = OSG (allc)4\*OSGA (allc)4

|  |  |
| --- | --- |
| \* | (Som (j) (OSLG(a1j1)\*OSLGA(a1j1c)))2 |
| (Som (j)(1/(OSLG(a1j1)\*OSLGA(a1j1c)))2 |

RDL = (OLG (a11) \* OLGA(a11e) \* OLG(a22) \* OLGA(a22c))2

(La catégorie 2 de la variable L indique la catégorie de référence pour cette variable, alors que la catégorie 1 varie de catégorie en catégorie de L.)

[221]

b) G est polytomique

RDS = (OSG (a11) \* OSGA (a11e) \* OSG (a22) \* OSGA(a22c))2

|  |  |
| --- | --- |
| \* | (Som (j) (OSLG (al j1)\*OSLGA (a1j1c))) |
| (Som (j) (OSLG (a1j2)\*OSGLA (a1j2c))) |
| (Som (j) (l/OSLG (a1j1)\*OSLGA (a1j1c))) |
| (Som (j) (l/(OSLG(alj2)\*OSLGA(alj2c))) |

|  |  |
| --- | --- |
| RDL = | OLG (a11) \* OLGA (a11e) \* OLG (a22) \* OLGA (a22c) |
| OLG (a12) \* OLGA (a12c) \* OLG (a21) \* OLGA (a21c) |

|  |  |
| --- | --- |
| \* | (Som (i) (OSGL (ai11)\*OSLGA (ai11e))) |
| (Som (i) (OSLG (ai1l2)\*OSLGA (ai12c))) |
| (Som (i) (OSLG (ai21)\*OSLGA (ai21c))) |
| (Som (i) (OSLG (ai22)\*OSLGA (ai22c))) |

Cas 3 : *S est polytomique et L est dichotomique*:

a) G est dichotomique

RDS = (OSG (a11) \* OSGA (a11e) \* OSG (a22) \* OSGA (a22c))2

RDL = (OLG (a11)4\*OLGA (a11c)4

|  |  |
| --- | --- |
| \* | (Som (i) (OSLG (ai11)\*OSLGA (ai11e)))2 |
| (Som (i) (l/(OSLG(ai11)\*OSLGA(ai11e)))2 |

La catégorie 2 de la variable S indique la catégorie de référence pour cette variable, alors que la catégorie 1 varie de catégorie en catégorie de S.

[222]

b) G est polytomique

|  |  |
| --- | --- |
| RDS = | OSG(a11) \* OSGA(a11e) \* OSG(a22) \* OSGA(a22c) |
| OSG(a12) \* OSGA(al2c) \* OSG(a21) \* OSGA(a21c) |

|  |  |
| --- | --- |
| \* | (Som(j)(OSLG(a1j1)\*OSLGA(a1j1e))) |
| (Som(j)(OSLG(a1j2)\*OSLGA(a1j2c))) |
| (Som(j)(OSLG(a1j2)\*OSLGA(a2j1c))) |
| (Som(j)(OSLG(a2j2)\*OSLGA(a2j2c))) |

RDL = (OLG(a11) \* OLGA(a11c) \* OLG(a22) \* OLGA(a22c))2

|  |  |
| --- | --- |
| \* | (Som(i)(OSLG(ai11)\*OSLGA(ai11e))) |
| (Som(i)(OSLG(ai12)\*OSLGA(ai12c))) |
| (Som(i)(l/OSLG(ai1l)\*OSLGA(ai11c))) |
| (Som(i)(l/(OSLG(ai12)\*OSLGA(ai12c))) |

Cas 4 : *S et L sont polytomiques*:

a) G est dichotomique

RDS = (OSG(a11) \* OSGA(a11c) \* OSG(a22) \* OLGA(a22c)2

|  |  |
| --- | --- |
| \* | (Som(j)(OSLG(a1j1)\*OSLGA(a1j1c))) |
| (Som(j)(l/(OSLG(a1j1)\*OSLGA(a1j1e))) |
| (Som(j)(OSLG(a2j1)\*OSLGA(a2j1c))) |
| (Som(j)(1/(OSLG(a2j1)\*OSLGA(a2j1c))) |

RDL = (OLG(all) \* OLGA(al le) \* OLG(a22) \* OLGA(a22c)2

|  |  |
| --- | --- |
| \* | (Som(i)(OSLG(ai11) \*OSLGA(ai11e))) |
| (Som(i)\*1/(OSLG(ai11)\*OSLGA(ai11c))) |
| (Som(i)(OSLG(ai21)\*OSLGA(ai21c))) |
| (Som(i)1/(OSLG(ai21)\*OSLGA(ai21c))) |

[223]

Les catégories 2 des variables S et L jouent le rôle de la catégorie de référence alors que les catégories 1 de ces deux variables varient selon que l’on compare les différents autres sous-groupes définis par S et L à ces groupes de référence, b) G est polytomique

|  |  |
| --- | --- |
| RDS = | OSG(a11) \* OSGA(a11c) \* OSG(a22) \* OSGA(a22c) |
| OSG(a12) \* OSGA(a12c) \* OSG(a21) \* OSGA(a21c) |

|  |  |
| --- | --- |
| \* | (Som(j)(OSLG(a1j1)\*OSLGA(a1j1c))) |
| (Som(j)(OSLG(a1j2)\*OSLGA(a1j2c))) |
| (Som(j)(OSLG(a2j1)\*OSLGA(a2j1c))) |
| (Som(j)(OSLG(a2j2)\*OSLGA(a2j2c))) |

|  |  |
| --- | --- |
| RDL = | OLG(a21) \* OLGA(a11e) \* OLG(a22) \* OLGA(a22c) |
| OLG(a12) \* OLGA(a12c) \* OLG(a21) \* OLGA(a21c) |

|  |  |
| --- | --- |
| \* | (Som(i)(OSLG(ai11)\*OSLGA(ai11c))) |
| (Som(i)(OSLG(ai12)\*OSLGA(ai12c))) |
| (Som(i)(OSLG(ai21)\*OSLGA(ai21c))) |
| (Som(i)(OSLG(ai22)\*OSLGA(ai22c))) |

4. Conclusion

Si je pouvais poursuivre le développement des modèles proposés dans cette communication, j’explorerais sans doute au moins trois voies.

En premier lieu, le fait qu’un terme indiquant un effet d’inégalités ou de discrimination soit nécessaire pour rendre compte d’un ensemble de données ne signifie pas que tous les coefficients log-linéaires qui lui sont associés ou que tous les rapports dont il fait partie soient statistiquement différents de l’unité (ou de zéro dans une perspective log-linéaire logarithmique) : il serait sans doute utile de pouvoir tester la signification statistique des produits formés par les différents rapports d’inégalités ou de discrimination proposés. En nous inspirant [224] de ce que permet l’analyse de covariance structurelle avec des modèles comme LISREL, il serait sans doute encore plus fécond pour des fins de comparaison de pouvoir contraindre certains de ces coefficients à être égaux à une quantité donnée ou égaux entre eux.

Un moyen terme serait peut-être de pouvoir introduire dans les modèles l’équivalent des variables auxiliaires bien connues en régression (« dummy variables ») afin de pouvoir évaluer la pertinence de certains effets. Par exemple, dans les modèles de discrimination ayant servi d’exemple dans cette communication, la variable G est trichotomique : qu’il soit nécessaire de tenir compte d’un effet comme OSG ou OLG ne nous dit pas si c’est le rapport comparant les premières occupations des répondants à celles de leurs pères, celui comparant les occupations des répondants à 35 ans à leurs premières ou, enfin, celui comparant les occupations à 35 ans à celles des pères qui est significatif : les trois peuvent l’être comme deux ou un seul. Une procédure claire permettant de construire plusieurs variables plus simples selon les états de G comparés permettrait de répondre à ce genre de questions. Pour l’instant, je ne vois que dans la définition de contrastes particuliers une solution à ce problème.

En troisième lieu, les rapports d’inégalités et de discrimination proposés ont l’avantage d’attirer « dramatiquement » l’attention sur les catégories professionnelles dans lesquelles les inégalités et/ou la discrimination sont les plus fortes. Ils ne donnent cependant aucune idée de l’importance relative de ces lieux d’inégalités ou de discrimination dans l’ensemble de la structure professionnelle comme le font les indices de dissimilarités de Duncan ou de Cortese (Duncan et Duncan, 1955, Cortese, Falk et Cohen, 1976) : il serait utile de pouvoir compléter l’analyse des rapports d’inégalités et/ou de discrimination par des tableaux « standardisés » ou purifiés comme le proposent Clogg et Eliason (1988) permettant de définir des indices de dissimilarités similaires à ceux évoqués.

De façon plus générale, il faut être conscient des limites de ces modèles. Réduit à son essence, le modèle des inégalités permet d’examiner l’influence d’une variable indépendante (par exemple, le sexe) sur une variable dépendante (ici, l’occupation) en contrôlant sur deux variables pouvant affecter cette relation (la « langue » et [225] l’année). Le modèle de la discrimination est encore plus « simpliste » puisqu’il permet de tenir compte explicitement de l’influence d’une seule variable pouvant « expliquer les inégalités : l’occupation du père (ou une occupation antérieure) (en plus de la langue et de l’année comme pour les inégalités). En d’autres termes, ce dernier modèle suppose que ce qui n’est pas expliqué par l’occupation du père (ou la première occupation) relève de la discrimination et non d’une autre source de variation « naturelle » comme la scolarité, le milieu socio- géographique d’origine, etc. Formellement, cette difficulté peut sembler assez facile à résoudre puisqu’il « suffirait » d’introduire dans les modèles log-linéaires de telles variables supplémentaires. Je ne suis pas certain cependant qu’il sera possible de définir alors des rapports d’inégalités ou de discrimination aussi simples que ceux proposés dans cette communication ou que la taille des échantillons qu’il faudra considérer pour tenir compte de tous les « effets » nécessaires ne devienne prohibitive dans les conditions actuelles de recherche.

[226]

RÉFÉRENCES BIBLIOGRAPHIQUES

F. BÉLAND et M. de SÈVE, L’inégalité et la discrimination sexuelles et linguistiques au Québec, *Revue Canadienne de sociologie et d’anthropologie,* 23 (3), 1986, pp. 309-330.

P. BERNARD*,* A. DEMERS, D. GRENIERet J. Renaud, *L’évolution de la* si*tuation des francophones et des non-francophones au Québec (1971-1978)*, Gouv. du Québec, Office de la langue française, 1980, 3ème tirage.

C.C. CLOGG et S.R. ELIASON, A Flexible Procedure for Adjusting Rates and Proportions, including Statistical Methods for Group Comparisons, *American Sociological Review*, vol. 53, 2, avril 1988, pp. 267-283.

CORTESE C.F., FALK R.F. et COHEN J.K., Further Considérations on the Methodological Analysis of Segregation Indices, *American Sociological Review*, 41, 4, 1976, pp. 630-637.

DUNCAN O.D. et DUNCAN B., A methodological analysis of segregation indexes, *American Sociological Review,* 20, 2, 1955, pp. 210-217.

HAGENAARS J. A., *Categorical Longitudinal Data, Log-linear Panel, Trend and Cohort Analysis*, Sage, 1990.

SANTI L.L., Partialling and Purging : Equivalences Between Log-Linear Analysis and the Purging Method of Rate Adjustement, *Sociological Methods and Research*, vol. 17, 4, mai 1989, pp. 376-397.

[227]

Annexe :  
Les rapports multiplicatifs purifiés

1. Les inégalités

Dans cette annexe, nous avons conservé la même notation pour indiquer les coefficients mesurant les inégalités : RIS et RIL. Comme l’indique la section 2.4, dans le cas des coefficients multiplicatifs, ces deux expressions indiquent alors le produit de plusieurs rapports (et non le rapport de deux sommes). Dans le cas des inégalités selon la variable S, les rapports femmes/hommes de tous les différents groupes selon la variable L (l’origine ethnique ou linguistique) sont multipliés :

|  |  |  |  |
| --- | --- | --- | --- |
| F(ac) | = P(j,1,n) | F(ajc) | où j varie de 1 à *n* selon le nombre de catégories de L |
| H(ac) | H(ajc) |

Dans le cas des inégalités selon la variable L (ethniques ou linguistiques), les rapports francophones/anglophones des différents groupes selon S sont multipliés :

|  |  |  |  |
| --- | --- | --- | --- |
| FR(ac) | = P(i,1,m) | Fr(aic) | où i varie de 1 à *m* selon le nombre de catégories de S |
| An(ac) | An(aic) |

Dans les formules suivantes, en particulier quand S ou L sont polytomiques, le groupe « 1 » indique le groupe comparé (par exemple, les femmes pour les inégalités sexuelles et les francophones pour les inégalités linguistiques) et le groupe « 2 », le groupe de référence (les hommes pour les inégalités sexuelles et les anglophones pour les inégalités linguistiques).

1.1. *S et L sont dichotomiques*:

RIS = (OS (ai))4 \* (OSA(a1c))4

RIL = (OL (a1))4 \* (OLA(a1c))4

[228]

1.2. *S est dichotomique et L est polytomique (n catégories)* :

RIS = (OS (a1))2n \* (OSA(a1c))2n

|  |  |
| --- | --- |
| RIL = | (OL (a1))2 \* (OLA(a1c))2 |
| (OL (a2))2 \* (OLA(a2c))2 |

1.3. *S est polytomique (m catégories) et L est dichotomique :*

|  |  |
| --- | --- |
| RIS = | (OS(a1))2 \* (OSA(a1c))2 |
| (OS(a2))2 \* (OSA(a2c))2 |

RIL = (OL(a1))2m \* (OLA(a1c))2m

1.4. *S et L sont polytomiques (m et n catégories) :*

|  |  |
| --- | --- |
| RIS = | (OS(a1))n \* (OSA(a1c))n |
| (OS(a2))n \* (OSA(a2c))n |

|  |  |
| --- | --- |
| RIL = | (OL(a1))m \* (OLA(a1c))m |
| (OL(a2))m \* (OLA(a2c))m |

2. La discrimination

L’équivalent du RDS est le produit sur les différents groupes ethniques j des rapports suivants :

|  |
| --- |
| OSG(a11) \* OSGA(a11c) \* OSLG(a1j1) \* OSLGA(a1j1c) |
| OSG(a12) \* OSGA(a12c) \* OSLG(a1j2) \* OSLGA(a1j2c) |
| OSG(a21) \* OSGA(a21c) \* OSLG(a2j1) \* OSLGA(a2j1c) |
| OSG(a22) \* OSGA(a22c) \* OSLG(a2j2) \* OSLGA(a2j2c) |

Le produit sur j permet de faire disparaître tous les termes OSLG et OSLGA car celui-ci égale 1 ; ainsi, si la variable Lan valeurs, le produit se réduit à l’expression suivante :

[229]

|  |
| --- |
| (OSG(a11) \* OSGA(a11e) \* OSG(a22) \* OSGA(a22c))n |
| (0SG(a12) \* 0SGA(a12c) \* OSG(a21) \* 0SGA(a21c))n |

De façon similaire, l'équivalent du rapport RDL se simplifie de la façon suivante (où m est le nombre de valeurs de la variable S):

|  |
| --- |
| (OLG(a11) \* OLGA(a11c) \* OLG(a22) \* OLGA(a22c))m |
| (OLG(a12) \* OLGA(a12c) \* OLG(a21) \* OLGA(a21c))m |

Selon que G est dichotomique ou non, nous pouvons distinguer 16 cas. (Dans tous ceux-ci, la valeur 1 est celle de la catégorie étudiée et la valeur 2, celle de la catégorie de référence).

2.1. *S et L sont dichotomiques*

2.1.a) G est dichotomique

RDS = (OSG(a11) \* OSGA(a11e))8

RDL = (OLG(a11) \* OLGA(a11e))8

2.1.b) G n'est pas dichotomique

RDS = (OSG(a21) \* OSGA(a11e) \* OSG(a22) \* OSGA(a22c))4

RDL = (OLG(a11) \* OLGA(a11c) \* OLG(a22) \* OLGA(a22c))4

2.2. *S est dichotomique et L est polytomique (n)*

2.2.a) G est dichotomique

RDS = (OSG(a11) \* OSGA(a11c))4n

RDL = (OLG(a11) \* OLGA(a11c) \* OLG(a22) \* OLGA(a22c))4

2.2.b) G n'est pas dichotomique

RDS = (OSG(a11) \* (OSGA(a11e) \* OSG(a22) \* OSGA(a22c))2n

|  |  |
| --- | --- |
| RDL= | (OLG(a11) \* OLGA(a11c) \* OLG(a22) \* OLGA(a22c))2 |
| (OLG(a12) \* OLGA(a12c) \* OLG(a21) \* OLGA(a21c))2 |

2.3. *S est polytomique (m) et L est dichotomique*

2.3.a) G est dichotomique

RDS = (OSG(a11) \* OSGA(a11e) \* OSG(a22) \* OSGA(a22c))4

RDL = (OLG(a11) \* OLGA(a11c))4m

[230]

2.3.b) G est polytomique

|  |  |
| --- | --- |
| RDS = | (OSG(a11) \* OSGA(a11c) \* OSG(a22) \* OSGA(a22c))2 |
| (OSG(a12) \* OSGA(a12c) \* OSG(a21) \* OSGA(a21c))2 |

RDL = (OLG(a11) \* (OLGA(a11e) \* OLG(a22) \* OLGA(a22c))2m

2.4. *S et L sont polytomiques (m et n)*

2.4.a) G est dichotomique

RDS = (OSG(a11) \* OSGA(a11c) \* OSG(a22) \* OSGA(a22c))2n

RDL = (OLG(al 1) \* (OLGA(a22c) \* OLG(a22) \* OLGA(a22c))2m

2.4.b) G est polytomique

|  |  |
| --- | --- |
| RDS = | (OSG(a11) \* OSGA(a11c) \* OSG(a22) \* OSGA(a22c))n |
| (OSG(a12) \* OSGA(a12c) \* OSG(a21) \* OSGA(a21c))n |

|  |  |
| --- | --- |
| RDL = | (OLG(a11) \* OLGA(a11c) \* OLG(a22) \* OLGA(a22c))m |
| (OLG(a12) \* OLGA(a12c) \* OLG(a21) \* OLGA(a21c))m |