

Modéliser l'évolution de l'inégalité des taux de scolarisation

Guy Morel

▶ **To cite this version:**

Guy Morel. Modéliser l'évolution de l'inégalité des taux de scolarisation. Mathémathiques et Sciences humaines, 2003, 41 (164), p. 5-27. halshs-00009014

HAL Id: halshs-00009014
https://shs.hal.science/halshs-00009014

Submitted on 10 Feb 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Math. & Sci. hum. / Mathematics and Social Sciences (41e année, n° 164, *2003*, p. 5-27)

MODELISER L'EVOLUTION DE L'INEGALITE DES TAUX DE SCOLARISATION [1]

Guy MOREL[2]

RESUME – *La modélisation de l'évolution des inégalités liées aux taux de scolarisation a suscité un long débat dans les années 84-97. Après avoir essayé de synthétiser les principales contributions nous considérons une modélisation basée sur une variable latente appelée « revenu scolaire ». Les taux de scolarisation à un niveau donné nous servent à estimer les moyennes de cette variable sur les catégories sociales concernées. Ce passage d'une propriété locale à une propriété plus globale peut s'interpréter comme un changement d'échelle de mesure. Beaucoup des échelles déjà proposées, pour juger des différences de taux, sont peu compatibles avec ce point de vue, mais ce n'est pas le cas de la plus utilisée actuellement : l'échelle logistique. Sous cet éclairage, elle n'est cependant pas un modèle idéal.*

MOTS-CLES – *Inégalités de scolarisation, Variable latente, Modèle logit*

SUMMARY – Models for the Evolution of Inequality in Educational Level

During the years 84-97, a long debate occurred in France about models for the evolution of inequalities related to the percentages of a given educational level. First we try to synthesize the main contributions, then we consider models based on a latent variable called : "educational income". The frequencies at a given level of education are used to estimate the means of this variable for the social categories concerned. This passage from a local property of a distribution to a more central property can be interpreted as a measurement scale. The scales comparing percentages, already suggested in the debate, are often not very compatible with this point of view. It is not the case of the currently most used one: the logistic scale. Under this light, it is not an ideal model however.

KEYWORDS – *Educational Level Inequalities, Latent Variable, Logit Model*

Comparer l'évolution de plusieurs proportions est un problème d'apparence simple. Le débat enclenché par Jean-Claude Combessie [1984] en montre pourtant toute la difficulté. Les nombreux articles parus dans la *Revue française de sociologie* et dans *Mathématiques et Sciences humaines* n'ont pas épuisé le sujet. Ils ont éclairé d'une lumière différente les multiples facettes du débat. Nous partons du même type d'exemple, celui de l'évolution de l'inégalité des chances devant l'enseignement, car la modélisation de l'évolution de proportions est ici tributaire du champ d'application. Il ne faut pas simplement choisir une bonne échelle de mesure pour des proportions mais aussi, bien traduire la notion de constance des inégalités.

---

[1] Article reçu le 31 décembre 2002, révisé le 25 mars 2003, accepté le 31 mars 2003.

[2] Département de sociologie, Université François Rabelais, 3 rue des Tanneurs, 37041 Tours Cedex.
Laboratoires : DSU-VST et CNRS UPRES-A 6083, e-mail : morel@univ-tours.fr

Nous proposons dans cet article un angle d'attaque différent en faisant intervenir une variable latente que nous appelons «□evenu scolaire□. Un taux de scolarisation à un niveau donné fournit une propriété de la distribution de cette variable sur la population concernée. Il n'en définit pas pour autant un bon indice pour l'étude de l'évolution de l'inégalité. Une bonne partie du débat qui nous intéresse porte d'ailleurs sur les transformations à faire sur ces taux pour pouvoir juger de l'évolution des inégalités. Nous allons fonder notre jugement sur l'évolution des revenus scolaires moyens. Les taux de scolarisation nous serviront à les estimer. Ce point de vue conduit à de nouveaux traitements, il permet aussi de retrouver des solutions déjà proposées. Certaines de ces solutions passent assez bien sous ce nouvel éclairage, d'autres moins.

Avant d'exposer notre point de vue nous devons essayer de synthétiser le débat débuté par Jean-Claude Combessie en 1984 et qui s'est prolongé pendant de nombreuses années. Dans *Mathématiques et Sciences humaines*, des articles récents y font encore référence (cf. par exemple [É. Terouanne, 1995] et [A. Valeyre, 1995]). Il en est de même dans la *Revue française de sociologie* (cf. par exemple [M. Euriat, C. Thelot, 1995]), le débat y semble cependant avoir été tranché, au niveau de l'échelle de mesure, en faveur du taux logistique à partir de l'article de Louis-André Vallet [1988]. Ces dernières années les articles traitant des inégalités sociales devant l'école utilisent souvent la régression logistique et les modèles log-linéaires. Il ne nous semble pas que les bonnes propriétés de ces outils, tant pour l'interprétation que pour le travail mathématique[3], mettent un point final au débat sur la modélisation statistique de l'évolution de l'inégalité des chances devant l'enseignement.

1. DIFFERENTS POINTS DE VUE

À la base du débat il y a l'exemple utilisé par Raymond Boudon dans *L'inégalité des chances* [1973] et construit à partir d'un rapport de John Westergaard et Alan Little [1967]. Il donne pour deux tranches d'âges et deux classes extrêmes les taux de scolarisation dans le secondaire long des enfants anglais nés dans la première moitié du siècle.

Catégories sociaux-professionnelles du père	Date de naissance	
	Avant 1910	Entre 1935 et 1940
Professions libérales, cadre et personnel de direction	37 %	62 %
Ouvriers semi-qualifiés et non qualifiés	1 %	10 %

Nous ne nous occuperons pas dans cet article des problèmes de statistique inductive : tests, intervalles de confiance. Ce tableau sera analysé comme s'il était issu de données exhaustives ou d'un échantillon suffisamment grand pour que les fluctuations d'échantillonnage aient de grandes chances d'avoir des effets négligeables sur l'interprétation sociologique. Cette position est adoptée dans la plupart des articles portés au débat. La critique des sources statistiques ayant conduit à l'exemple traité a été effectuée par Jean-Paul Grémy [1984], bien que les taux analysés soient obtenus à partir

[3] Ils entrent dans le cadre des familles exponentielles (cf. par exemple [A. Antoniadis, J. Berruyer, R. Carmona, 1992], [R. Christensen, 1990], [X. Guyon, 2001]).

d'échantillons importants ils donnent des intervalles de confiance dont les longueurs ne sont pas totalement négligeables.

Cet exemple est repris dans plusieurs livres de méthodologie ou de statistique, car la comparaison de l'évolution des taux de scolarisation dans les deux classes conduit à un jugement différent suivant que cette évolution est mesurée en faisant la différence ou en faisant le rapport (62 % - 37 % > 10 % - 1% alors que 62 %/37 % < 10 %/1 %). Dans la première partie de son article, Jean-Claude Combessie [1984, p. 235] passe en revue les opérations élémentaires que l'on peut faire sur ces quatre pourcentages pour porter un jugement. Il en décrit 24, plusieurs étant équivalentes. Dominique Merllié [1985, p. 642] montre qu'il y a trois manières fondamentales de procéder, les deux décrites précédemment (différence et rapport) et le rapport des taux de non scolarisation (rapport des pourcentages complémentaires). Le rapport des taux de non scolarisation indique une diminution moins rapide chez les ouvriers (90 %/99 % > 38 %/63 %) alors que les rapports des taux de scolarisation nous ont conduits à un jugement différent puisque l'augmentation est moins rapide chez les cadres. Cette structure des concordances et des oppositions entre les conclusions produites par ces trois regards n'est pas constante (cf. [D. Merllié, 1985, p. 639]), elle dépend des données.

1.1. CHOIX D'UNE ÉCHELLE DE MESURE

L'échelle des différences conduit à dire que l'écart entre 4 % et 8 % est équivalent à l'écart entre 20 % et 24 %, alors que l'échelle de rapport rend le premier écart équivalent à celui entre 20 % et 40 % (multiplication par 2). Ceci revient à utiliser l'échelle logarithmique c'est-à-dire à prendre pour écart entre deux taux la différence de leur logarithme (ln(40 %) - ln(20 %) = ln(40/20) = ln(2) = ln(2p) - ln(p)). Pour le rapport des taux de non scolarisation (1 - p) le passage de (100 - 20) % à (100 - 40) % est équivalent au passage de (100 - 4) % à (100 - 28) %[4], c'est considérer l'échelle logarithmique suivante : - ln(1 - p).

Définir une échelle de mesure revient à se donner une fonction *f* continue strictement croissante[5]. Cette fonction définit «une dimension implicite, ou latente, servant à évaluer la signification (ou l'importance) que le chercheur attribue à ces observations» [J.-P. Grémy, 1984, p. 419]. Remarquons que la comparaison des évolutions dans les deux classes (f(62 %) - f(37 %) comparé à f(10 %) - f(1 %)) donne le même résultat que l'étude de l'évolution de l'écart entre classes (passage de f(37 %) - f(1 %) à f(62 %) - f(10 %))[6]. La vision diachronique et la vision synchronique donnent la même chose, il n'en sera pas ainsi dans la généralisation à plus de deux classes et/ou à plus de deux générations, nous le verrons plus loin. Continuons d'abord l'étude du choix d'une échelle *f*. Un changement d'origine et d'unité de l'échelle ne change pas l'interprétation des comparaisons, f et $af + b$ ($a > 0$) donnent le même type de conclusions. Cependant un grand nombre d'échelles de mesure restent possibles. On peut se demander si certaines ne sont pas inadaptées au problème.

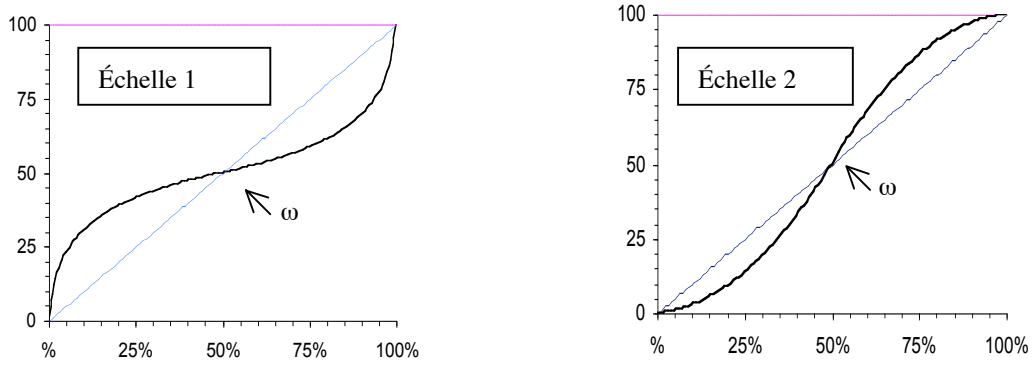
La contradiction possible entre les deux échelles de rapport est particulièrement choquante, elle conduit Jean Prévot [1985, p. 605] à imposer le principe de dualité «

[4] (100 - x)/(100 - 4) = (100 - 40)/(100 - 20) ⇔ 100 - x = 96 * 3/4 = 72.

[5] Fonction strictement croissante pour conserver l'ordre induit par les pourcentages (cf. [J. Prévot, 1985, p. 613]), continue pour qu'un écart infinitésimal de pourcentages ne puisse pas induire un saut sur l'échelle de mesure.

[6] a - b > c - d ⇔ a - c > b - d.

paraît naturel et raisonnable de soumettre à ce principe tout indice relatif à l'évolution de *proportions dichotomiques* (ici taux d'admis et taux d'exclus), à savoir qu'appliqué à l'*évolution des proportions complémentaires* (les taux d'exclus) cet indice doit conduire à la «□même mesure□ (i.e. même sens et même intensité) que lorsqu'on l'applique aux proportions premières (les taux d'admis)□. Pour une échelle de mesure ceci se traduit par : $f(p_2) - f(p_1) = f(1 - p_1) - f(1 - p_2)$[7], ce qui est équivalent à la symétrie suivante : $[f(p) + f(1 - p)]/2 = f(50\ \%)$[8]. Si f est bornée on pose souvent $f(0) = 0$ et $f(100\ \%) = 100$[9] donc $f(50\ \%) = 50$, dans le cas contraire on imposera plutôt $f(50\ \%) = 0$. Les Figures ci-dessous donnent deux exemples d'échelles bornées vérifiant le principe de dualité. Dans chaque cas, les deux courbes sont symétriques par rapport au centre ω du carré et la diagonale représente l'échelle des différences, elle vérifie le principe de dualité et conserve les écarts entre pourcentages. L'échelle 2, elle, diminue l'écart entre deux pourcentages quand ils sont proches de 0 % ou de 100 %. C'est plutôt l'inverse (l'augmentation) qui est recherchée, ce que fait l'échelle 1.



Graphique 1

Parmi les échelles non bornées les deux échelles de rapport ne vérifient pas le principe de dualité. Une autre échelle non bornée, couramment utilisée en statistique dans le maniement des pourcentages est le taux logistique : $f(p) = \ln(p/(1 - p))$. Cette échelle basée sur le rapport des chances d'être scolarisé plutôt que non scolarisé $(p/(1 - p))$ vérifie le principe de dualité. De plus elle se prête bien à une interprétation sociologique (cf. par exemple [L.-A. Vallet, 1988]). Cette dernière propriété est fondamentale pour choisir entre les différentes échelles vérifiant le principe de dualité. D'ailleurs Louis-André Vallet [1988] ne retient comme échelles que celle des différences (les pourcentages observés) et le taux logistique. Sur l'exemple de départ on obtient avec cette échelle une augmentation plus rapide chez les ouvriers $(f(10\ \%) - f(1\ \%) = \ln(10/90) - \ln(1/99) = 2.4)$ que chez les cadres $(f(62\ \%) - f(37\ \%) = \ln(62/38) - \ln(37/63) = 1.0)$.

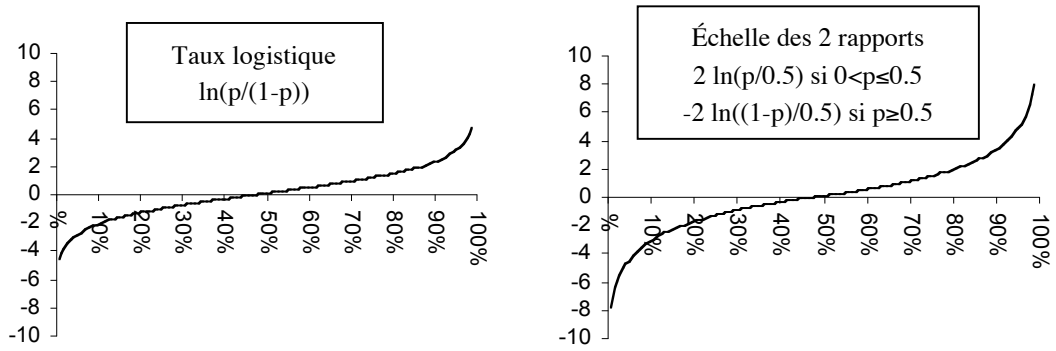
À ce point de l'exposé de nombreuses autres échelles sont possibles mais difficilement justifiables. On pourrait tout au plus proposer l'échelle obtenue en raccordant les échelles de rapport au point 50 %. L'échelle ln(p) entre 0 % et 50 % puisqu'elle a été construite pour que l'écart entre deux petits pourcentages x % et 2x %

[7] Au niveau de l'interprétation, la croissance du taux d'admis correspond à la décroissance du taux d'exclus.

[8] Il suffit d'utiliser l'égalité : $f(50\ \%) = f(100\ \% - 50\ \%)$.

[9] Ceci est toujours possible par transformation linéaire.

soit équivalent à l'écart entre 20 % et 40 %. Cet écart sera aussi équivalent à celui entre 92 % et 96 % avec l'échelle - ln(1 - p)[10]. Le Graphique 2 permet de comparer cette échelle des 2 rapports avec celle du taux logistique. L'échelle des 2 rapports différencie encore plus un écart entre deux pourcentages proches de 0 % (ou de100 %) d'un même écart autour de 50 %.



Graphique 2

Nous reviendrons sur le choix d'une échelle avec le nouvel éclairage que nous allons introduire. Un tel choix ne résout cependant pas tous les problèmes posés dans le débat qui a eu lieu autour de cet exemple simple. Une question est particulièrement importante : est-il légitime de comparer des taux d'admis au même niveau scolaire à des périodes très différentes du point de vue du niveau global de scolarisation ? L'interprétation de l'évolution des taux de scolarisation ne se ramène pas au problème de l'évolution de pourcentages, il y a une difficulté supplémentaire due au fait que les modalités de la variable niveau scolaire sont nominalement les mêmes alors que leur significations sociales varient dans le temps.

1.2. ÉVOLUTION À INÉGALITÉ CONSTANTE

L'évolution globale de la scolarisation influe sur le taux de scolarisation des deux classes. Pour savoir si les différences observées sont dues à une évolution qui diminue, conserve ou augmente les inégalités, il faut essayer de définir la notion de constance des inégalités dans un système scolaire en mouvement. Le taux de scolarisation à un niveau scolaire donné est un indicateur qui n'a pas la même signification aux différentes dates. Dominique Merllié [1985, p. 647] prend une comparaison parlante :

> *On se trouve dans la même situation que si on voulait juger de l'évolution de l'écart des revenus de groupes sociaux en connaissant seulement la proportion de ceux qui dépassent à chaque date un revenu d'un montant constant en francs courants : si l'on passe d'une situation où 1 % des ouvriers ont un salaire d'au moins tant par mois contre 37 % des cadres à une situation où la même somme est gagnée par 10 % des ouvriers et 62 % des cadres, comment juger si la tendance a été à un resserrement de la hiérarchie des salaires ? Nul doute que ceux-là mêmes qui concluent à une évolution « incontestable » en matière de scolarisation ne s'interdisent dans*

[10] Pour avoir f(50 %) = 0, on pose f(p) = ln(p) - ln(0.5) sur]0 ; 50 %] et f(p) = ln(0.5) - ln(1 - p) sur [50 % ; 100 %[. On a multiplié par deux cette fonction afin d'avoir pour p = 50 % la même tangente que le taux logistique.

> *ce cas, pourtant formellement identique, à conclure. Pour y voir plus clair – et en admettant qu'on doive se contenter d'un indicateur aussi pauvre que le seul taux de ceux qui atteignent ou dépassent un montant donné –, il faudrait au moins que ce montant ne soit pas « le même » aux deux dates, mais qu'on sache par exemple quelle est, à la deuxième date, la proportion d'ouvriers dont le salaire dépasse celui qui est dépassé par 37 % des cadres. C'est en renonçant à tenir (nominalement) constant l'indicateur retenu qu'on se donnerait un point structurellement fixe de comparaison, permettant de juger si, par rapport à lui, la situation des ouvriers s'est plutôt dégradée ou améliorée.*

Un principe semblable est appliqué pour la définition de la pauvreté, il conduit à la notion de seuil relatif de pauvreté.

L'application de cette remarque conduit à comparer des taux de scolarisation à des niveaux différents jugés équivalents. Dans son exemple illustratif, Dominique Merllié [1985, p. 649] ne compare pas les taux d'accès en 6e pour l'année 1962, avec ceux des années 1972-73-74 mais avec les taux d'accès dans les seconds cycles de l'enseignement secondaire[11]. Il justifie cette comparaison de taux d'accès à des niveaux différents, par le fait que ces niveaux sont dans les générations concernées globalement atteint par le même pourcentage d'enfants. Cet argument est aussi utilisé par Michel Euriat et Claude Thélot [1995, p. 414] pour justifier la comparaison des recrutements de quatre grandes écoles entre 1950 et 1990. Le nombre de places offertes a augmenté mais il représente toujours environ 1,2 ‰ d'une génération. Les auteurs parlent de « comparabilité maximale ». La fréquence globale d'un niveau scolaire est donc pris comme indicateur de sa « valeur intrinsèque » à la date considérée. Il n'y a bien sûr pas qu'une seule manière possible de définir cette « valeur intrinsèque ». Une fois cette définition choisie il y a la difficulté technique de trouver des statistiques comparables, les niveaux scolaires à un moment donné ne forment pas un continuum bien ordonné.

Certains auteurs préfèrent garder le même niveau de scolarisation – à condition qu'il ne soit pas devenu trop généralisé – en essayant de modéliser la notion d'évolution à inégalité constante. Ils tentent de répondre à la question suivante : si la première génération avait eu un taux global de scolarisation, au niveau scolaire étudié, égal à celui de la seconde génération, qu'aurait-il fallu observer comme taux de scolarisation dans les différentes classes pour garder la même inégalité ? Une fois cette modélisation choisie il est alors possible de comparer les taux reconstitués de la première génération à ceux observés dans la seconde. Ce choix de modélisation est souvent implicite il renvoie à l'idée « naturelle » qu'un changement scolaire simplement quantitatif (augmentation de la longueur de la scolarité) a un effet additif sur les mesures faites dans les différentes classes. Il n'est cependant pas complètement spécifié car il dépend de l'échelle de mesure dans laquelle se traduit l'additivité, ce qui nous renvoie au paragraphe précédent. Louis-André Vallet [1988] regarde s'il y a additivité sur les pourcentages observés et sur les taux logistiques. Cette dernière modélisation a sa préférence [L.-A. Vallet, 1988 et 1999]. La non interaction entre le facteur classe sociale et le facteur génération est pris comme traduction du non changement des inégalités. Cette traduction n'est pas indépendante de l'échelle de mesure choisie. Nous l'interpréterons en faisant intervenir une variable latente, l'équivalent de la variable

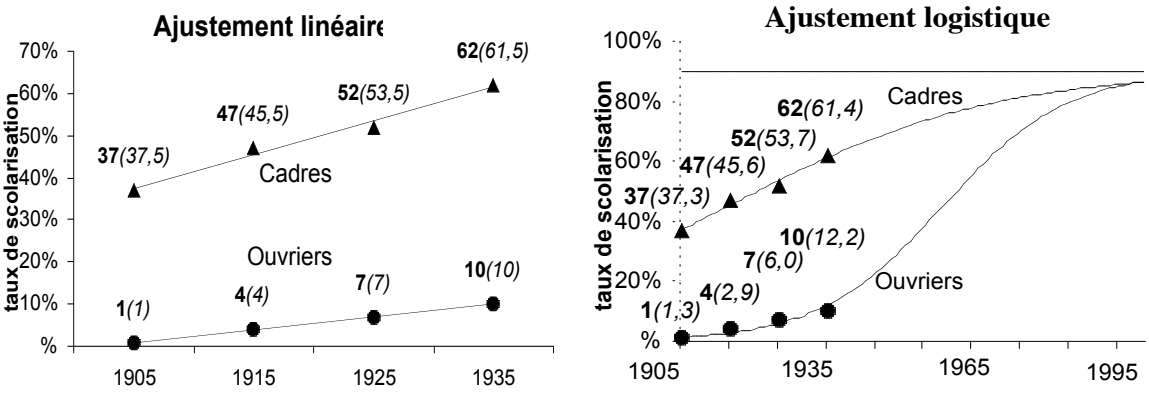
[11] Les taux d'accès en 6e en 1972-73-74 sont de plus peu pertinents pour la recherche d'un indice global d'inégalité, car tous proches de 100 %. Bien entendu, l'étude des non scolarisés en 6e montrerait sûrement une répartition sociale très inégale.

revenu dans la comparaison faite au début de ce paragraphe. Avant nous allons reprendre quelques problèmes soulevés dans le débat et liés à la généralisation aux cas de plus de deux classes et/ou de plus de deux générations.

Les deux grands types de solutions que nous venons de voir se prolongent à plusieurs C.S.P. et plusieurs cohortes. L'exemple illustratif de Dominique Merllié [1985] et celui de Louis-André Vallet [1988] portent d'ailleurs sur plus de deux C.S.P. Le choix du niveau scolaire de 1972-73-74 équivalent à celui du niveau de 6e en 1962 est valable pour toutes les C.S.P. Si on avait une troisième cohorte il faudrait trouver un nouveau niveau scolaire comparable aux deux précédents en appliquant le même procédé que celui utilisé pour définir l'équivalence entre les deux premiers niveaux. Quant à la traduction d'une constance de l'inégalité par l'additivité des mesures associées aux taux de scolarisation à un même niveau scolaire, elle s'étend à plusieurs générations et plusieurs C.S.P., comme dans le modèle d'analyse de variance à deux facteurs quand chacun des facteurs passe de deux à plusieurs modalités[12].

1.3. CAS DE PLUSIEURS COHORTES

Jean Prévot [1985, p. 615] propose un point de vue différent de ceux de Dominique Merllié [1985]) et de Louis-André Vallet [1988], il utilise des régressions temporelles. Pour chaque C.S.P. il s'intéresse à la courbe de croissance des taux de scolarisation à un même niveau. Les données du Graphique 3 complètent celles de l'exemple de départ en considérant deux générations intermédiaires[13].



Graphique 3
(les données sont en gras, les valeurs ajustées en italique)

Sur les données de chaque C.S.P. une régression linéaire semble légitime. La méthode des moindres carrés donne la droite d'ajustement de pente 0.8 pour les cadres et de pente 0.3 pour les ouvriers. Ces deux droites s'écartent, ce qui montre une divergence des taux de scolarisation et se traduit par une tendance à l'augmentation de «□inégalité des chances□. Dans cette présentation la constance de l'inégalité se traduit par l'égalité des pentes des deux droites, leur parallélisme. Dans le cas d'une bonne

[12] $m_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}$, $1 \le i \le k$, $1 \le j \le \ell$, dans le cas de l'additivité $\gamma_{ij} = 0$.

[13] Les «□générations moyennes□ sont considérées comme régulièrement espacées $t_1 = 1905$, $t_2 = 1915$, $t_3 = 1925$ et $t_4 = 1935$ (cf. la note 6 de Prévot [1985]).

approximation des données par les droites de régression, ce parallélisme des droites correspond à un modèle additif sur les taux de scolarisation. C'est un modèle additif particulier car la linéarité impose des augmentations proportionnelles à la durée[14]. Entre 1905 et 1915 l'augmentation du taux de scolarisation serait égale chez les cadres et les ouvriers mais elle serait aussi égale à l'augmentation entre 1915 et 1925 dans les deux classes. Cette méthode introduit une hypothèse sur le processus d'évolution du changement structurel affectant la scolarisation. Elle suppose une évolution «□néaire□ du système scolaire. Si les données vérifient cette hypothèse le seul avantage par rapport au modèle additif général est de synthétiser pour chaque classe les trois augmentations observées en un seul indice : la pente de la droite de régression.

Bien entendu les taux de scolarisation d'une C.S.P. peuvent évoluer d'une façon non linéaire. Jean Prévot [1985, p. 619] en donne deux exemples, l'ajustement exponentiel et l'ajustement logistique de la forme : p□=□exp(*a*t□+□b) et p = S/[1 + exp(-*a*t - b)]. L'ajustement logistique a l'avantage d'être une fonction variant entre 0 et le seuil S, il ne donne donc que des valeurs comprises entre 0 % et 100 % si S est inférieur ou égal à 100 %. Cette propriété est surtout importante dans le cadre d'une étude prévisionnelle.

La deuxième figure du Graphique 3 donne pour les cadres et les ouvriers l'ajustement logistique avec le seuil S = 90 %. Le cas S = 100 % est formellement équivalent à l'ajustement linéaire sur les logarithmes des rapports des chances p/(1 - p). La pente de la droite est le paramètre *a* de la fonction logistique 1/[1 + exp(-*a*t - b)]. Comme nous l'avons vu précédemment, cette pente résume les trois augmentations entre les quatre taux logistiques de la classe considérée[15]. Sur l'exemple, on trouve évidemment que l'augmentation chez les ouvriers est plus rapide que chez les cadres ($a_{\text{cadres}} = 0.033 < a_{\text{ouvriers}} = 0.078$). Cette conclusion, opposée à celle obtenue avec la régression linéaire, nous montre l'importance du choix de la forme du modèle de régression, même si entre 1905 et 1935 les modèles linéaire et logistique sont très proches. Les paramètres servant à définir un modèle renvoient à des propriétés globales qui ne sont pas limitées au domaine des valeurs observées utilisées pour estimer les paramètres.

Il y a une deuxième manière de lire les différences entre les deux courbes d'ajustement logistique du Graphique 3 (avec un seuil S de 100 % la forme de la figure est semblable). On peut conserver l'échelle des différences et porter un jugement sur la manière dont la distance entre les deux courbes évolue. Ces deux courbes commencent par s'écarter pour se rapprocher et se rejoindre ensuite. Localement, entre les années 1905 et 1935, on obtient la même conclusion qu'avec la régression linéaire, c'est celle obtenue précédemment à partir de l'échelle des différences. Dans une interprétation globale on obtiendra plutôt la conclusion énoncée, sans cette modélisation, par Leila Sussman [1969, p. 429] : «□l faudra atteindre une quasi-saturation de la demande d'éducation au niveau supérieur, émanant des classes sociales les plus élevées, avant qu'une nouvelle expansion ne réduise les écarts entre les taux de participation des différentes classes sociales□ (cf. [J. Prévot, 1985, p. 623]). Ce type de conclusion ne répond pas vraiment au problème qui nous préoccupe dans cet article : l'évolution de l'inégalité des chances devant l'enseignement. Le taux de scolarisation étudié ayant

[14] $m_{ij} = \mu + \alpha_i + \beta t_j$, $1 \le i \le 2$, $1 \le j \le 4$, au lieu de $m_{ij} = \mu + \alpha_i + \beta_j$.

[15] La dérivée de la fonction logistique p = 1/[1 + exp(-*a*t - b)] est égale à *a*p(1 - p). La croissance de p est donc d'autant plus forte que le paramètre *a* est grand.

tendance à s'uniformiser, il finit par perdre son aptitude à servir d'indicateur global de différenciation sociale. Ce niveau de scolarisation est devenu un bien courant, son étude met en avant des inégalités extrêmes qui sont plus difficilement interprétables en termes d'évolution globale des inégalités. Par contre ce processus de saturation nous montre bien la difficulté engendrée par le déplacement du champ d'expression de ce que nous nommons l'inégalité des chances devant l'enseignement.

Des remarques semblables peuvent être faites avec le modèle d'ajustement exponentiel. Il correspond lui à un ajustement linéaire sur les valeurs de l'échelle de rapport ln(p), mais on peut aussi lire l'évolution sur l'échelle des différences de taux. L'interprétation de ce type de modèle est plutôt local car sa forme n'en fait pas un bon modèle prévisionnel à long terme puisqu'il n'est pas borné.

Dans le problème que nous étudions, le choix d'un modèle d'ajustement ne nous semble pas devoir reposer sur la seule contrainte d'une bonne prévision. Dans une première étape, il nous semble important de choisir l'échelle de mesure qui permettra de transformer les taux de scolarisation en valeurs directement comparables pour juger des inégalités. Le choix d'un modèle d'ajustement sur ces mesures intervenant dans une deuxième étape. Si le modèle de régression doit bien s'ajuster aux données, il doit aussi permettre une comparaison facile avec le modèle choisi pour enlever les effets du changement de la structure de scolarisation sous l'hypothèse de stabilité des inégalités. Une lecture directe des courbes d'ajustement oublie la première étape. En regardant l'évolution de l'écart entre ces courbes, on opte pour l'échelle des différences. Sans y réfléchir, la stabilité des inégalités est traduite par une évolution «□arallèle□ des deux courbes. Cet écart constant est la traduction d'une augmentation identique des taux de scolarisation entre deux moments t_1 et t_2 quelles que soient les valeurs de ces taux en t_1. Si ces valeurs sont très différentes les potentialités de croissance ne sont pourtant pas les mêmes. Cet argument est souvent utilisé pour justifier un changement d'échelle, nous y reviendrons après avoir analysé les difficultés engendrées par la prise en compte de plus deux C.S.P.

1.4. CAS DE PLUSIEURS C.S.P.

Nous avons déjà vu que les solutions proposées par Dominique Merllié et par Louis-André Vallet restent valables dans le cas de plus de deux C.S.P. On peut aussi comparer plus de deux courbes d'ajustement. Pour toutes ces analyses, les évolutions comparées des taux de scolarisation dans les différentes C.S.P. peuvent conduire à des conclusions différentes suivant les comparaisons. Certaines peuvent s'interpréter comme une augmentation des inégalités alors que d'autres tendent plutôt à faire voir une diminution. Ceci n'est d'ailleurs pas un défaut puisque l'évolution des inégalités n'est pas forcément unidirectionnelle. Cependant certains auteurs ont essayé de proposer une réponse synthétique pour pouvoir donner une appréciation globale. Ils essayent de construire un indice global d'inégalité entre les diverses C.S.P. pour ensuite regarder son évolution.

La notion d'égalité ne pose pas problème, tous les auteurs considèrent que dans ce cas, la distribution des C.S.P. chez les admis au niveau scolaire choisi doit être identique à celle sur les exclus. Ceci revient à dire que les taux d'admission dans les différentes C.S.P. sont égaux ou que les variables C.S.P. et admission sont indépendantes. Une même définition de l'égalité peut cacher deux notions d'inégalité des taux de scolarisation : l'une s'intéresse aux inégalités entre classes en accordant le

même intérêt à chaque C.S.P. quel que soit son poids numérique, l'autre concerne l'inégalité des taux sur la partition engendrée par les C.S.P.. Dans le premier cas, seuls les taux de scolarisation dans les différentes C.S.P. sont pris en compte, alors que dans le second cas on utilise en plus la répartition des C.S.P. donc leur poids numérique. Les auteurs que nous avons cités jusqu'à maintenant n'utilisent que les taux de scolarisation, ils s'intéressent donc aux inégalités entre classes. Ce point de vue permet de simplifier le problème en ne prenant pas en compte la variation au cours du temps de la répartition des C.S.P. Pour deux taux de scolarisation fixés relatifs à deux classes C1 et C2, on ne se demande pas si l'inégalité est plus grande avec C1 représentant 10 % ou 80 % de la population. Dominique Merllié [1985, p. 646] le justifie ainsi «□ n'y a pas plus lieu de s'interdire de faire abstraction de la taille relative des groupes pour juger de la manière différente dont ils sont affectés par une variable, que de renoncer, d'une manière plus générale, aux différentes techniques de neutralisation des effets de structure□. Ces auteurs n'essayent d'ailleurs pas d'obtenir un indice global d'inégalité, ils font abstraction du poids des classes plutôt que de les considérer comme de même poids. S'ils voulaient utiliser un indice global d'inégalité ils devraient choisir un indice synthétisant les écarts entre les taux observés dans les différentes classes et un taux moyen correspondant à la situation d'égalité des taux. C'est un cas particulier de recherche d'un indice global dans le cadre de la deuxième notion d'inégalité□l'inégalité des taux sur la partition des C.S.P.. Simplement, les effectifs relatifs des différentes catégories pour chaque cohorte étudiée ne sont pas utilisés, on les remplace par une constante (l'inverse du nombre de C.S.P.). Ceci revient à imposer une répartition uniforme des C.S.P., c'est-à-dire considérer un tableau croisé (C.S.P. et modalités admis-exclus) ayant les taux de scolarisation observés et un même effectif pour chaque C.S.P.[16]

Regardons maintenant les solutions proposées pour l'étude de l'inégalité des taux, en prenant en compte la répartition des C.S.P..

Jean-Pierre Florens [1984] s'intéresse aux indices globaux d'inégalité basés sur une mesure de la dépendance. Il considère le tableau croisé des variables C.S.P. et admission (2 modalités admis-exclus) et cherche à construire un indice exprimant l'écart entre le tableau observé et le tableau théorique d'indépendance, qui correspond ici à la situation d'égalité. L'auteur considère la classe des mesures appelées les φ-divergences[17], elle contient en particulier la distance du khi-deux. Bien sûr ces indices globaux d'inégalité peuvent évoluer de façon contradictoire quand on passe d'une cohorte à une autre, l'un peut croître alors que l'autre décroît. Le sociologue se trouve devant un choix difficile.

Marc Barbut [1984(a),(b)] introduit, lui, un outil très utilisé par les économistes pour étudier les inégalités : les fonctions de concentration de C. Gini. Elles permettent une représentation graphique de la distorsion entre deux distributions définies sur un même ensemble de modalités, ici les C.S.P.. Nous avons trois distributions possibles suivant la population considérée : les admis, les exclus ou toute la cohorte. Notons ces

[16] On peut aussi imposer une répartition des C.S.P. «□roche□ de celles observées dans les différentes cohortes. Pour enlever l'effet de structure, il faut garder la même pour toutes les cohortes.

[17] Notons f_{ij} la fréquence observée pour la i□ C.S.P et la j□ modalité de scolarisation (j = 1 = admis, j = 2 = exclus), t_{ij} représente la fréquence théorique correspondant à l'indépendance. Si φ est une fonction réelle convexe telle que $\varphi(1) = 0$, la φ-divergence associée est définie par $\Sigma_{i,j}\, t_{ij}\, \varphi(f_{ij}/t_{ij})$. La distance du khi-deux s'obtient avec $\varphi(x) = (x-1)_$.

distributions D_{ad}, D_{ex} et D_{co}. On peut étudier trois types de courbes de concentration, suivant que l'on regarde la distorsion entre D_{ad} et D_{co} , entre D_{ex} et D_{co} ou entre D_{ad} et D_{ex}. Marc Barbut a étudié les deux premières courbes, il montre [M. Barbut, 1984(b)] qu'elles ne sont pas équivalentes pour étudier l'évolution de l'inégalité, l'une peut exprimer une augmentation de l'inégalité, l'autre une diminution. Ceci les rend critiquables du point de vue du principe de dualité énoncé dans le paragraphe sur le choix d'une échelle de mesure (cf. [J.-C. Combessie, 1985]). La distorsion entre les distribution D_{ad} et D_{ex} échappe à cette critique (cf. [É. Terouanne, 1995]). Ce principe de dualité entre souvent en contradiction avec les propriétés exigées d'un bon indicateur d'inégalité (cf. [M. Barbut, 1985, p. 16]), sous l'éclairage apporté par l'introduction de la variable latente «□revenu scolaire□ nous verrons qu'il implique un principe de symétrie trop fort.

La comparaison des courbes de concentration ne donne pas toujours le sens d'une évolution, ce n'est pas un indice quantitatif global d'inégalité mais une représentation plus riche qui peut donner à voir des évolutions non unidirectionnelles. On peut cependant construire un indice global d'inégalité à partir de ces courbes (le plus connu est l'indice de concentration de Gini). La justification du choix d'un indice reste un problème difficile. Comme le montre Marc Barbut [1985], il n'existe pas toujours d'indice vérifiant toutes les propriétés dont l'utilisateur aimerait le voir paré. D'autre part, l'évolution d'un indice d'inégalité globale ne peut s'analyser de façon isolée que si l'ensemble des inégalités qui le composent varient de façon homogène. Un indice peut rester constant alors qu'il y a eu un bouleversement social. Un tableau croisant C.S.P. et admis-exclus donne à voir une société qui peut être totalement différente si on permute les noms des C.S.P. sans permuter les données correspondantes, quel que soit l'indice choisi il restera pourtant constant. Les valeurs d'un indice ne sont pas les seules informations prises en compte pour l'interpréter, par exemple le sens d'une distance du khi-deux sur un tableau croisé passe souvent par la lecture des contributions de chaque case (cf. aussi [É. Terouanne, 1995, p. 37] pour la décomposition de l'indice de concentration de Gini).

La recherche d'un indice global d'inégalité nous semble ne pouvoir être que la dernière étape d'un processus de traduction rendant opératoire la problématique du sociologue. Nous allons commencer par nous demander ce qu'un taux de scolarisation est censé mesurer. La question se pose puisque différentes échelles de mesures sont proposées.

2. INEGALITE DE «□REVENU SCOLAIRE□

Nous allons ici utiliser la notion de «□revenu scolaire□ pour établir un parallèle avec un autre grand type d'inégalités étudiées : celles des revenus. Elle repose sur la prise en compte d'une variable latente qui quantifierait les différents niveaux de sortie du système scolaire. Cette modélisation est bien sûr simplificatrice, les différentes portes de sortie du système scolaire sont difficilement classables sur une seule dimension. La difficulté augmente avec la diversification des filières, le nombre d'années d'étude. Notre but n'est cependant pas de construire explicitement cette variable, mais de supposer son existence afin d'étudier les propriétés des mesures d'inégalités construites à partir des taux de scolarisation. Le taux de scolarisation d'une C.S.P., notée C_i, au niveau N et au temps T_j, est alors la fréquence des individus de C_i et de la cohorte

étudiée qui donnent à la variable latente une valeur supérieure au revenu scolaire $R_j(N)$ correspondant au niveau N. Bien que ce niveau scolaire soit fixe, la valeur du revenu scolaire correspondant peut changer avec le temps T_j. Un même diplôme ne représente pas la même chose à différentes époques. Ce taux de scolarisation est une propriété de la distribution de la variable latente : revenu scolaire, sur les individus de la cohorte étudiée appartenant à C_i. Ce n'est cependant pas une propriété qui correspond à un indice très utilisé dans l'étude des inégalités de revenus. La comparaison à partir d'un indice de tendance centrale, moyenne des revenus par exemple, semble plus adaptée pour résumer l'évolution des inégalités. Bien entendu ces indices supposent la connaissance de la distribution de la variable latente pour chacune des C.S.P.. Nous n'avons pas cette information, elle est même impossible à obtenir car la variable latente est difficile à définir explicitement. Pour avoir un ordre de grandeur de ces indices il nous faut faire des hypothèses sur la forme de la distribution de la variable latente[18]. Ces hypothèses sont d'autant plus restrictives que nous avons peu de données réelles sur ces distributions. Avec un simple taux de scolarisation à un niveau N, nous devrons nous restreindre à une famille de distributions ne dépendant que d'un seul paramètre réel. Le taux de scolarisation déterminera une seule distribution si la famille de distribution choisie, paramétrée par un nombre réel θ, donne des fonctions de répartition strictement ordonnées en θ[19]. On choisit généralement une paramétrisation conduisant à une décroissance stricte, ainsi plus θ est grand et plus la variable latente se distribue sur les niveaux élevés de revenus scolaires. De nombreuses familles de probabilités peuvent convenir. Nous allons en étudier quelques unes, en particulier parmi celles pour lesquelles θ correspond à la moyenne. Nous pourrons ainsi interpréter les échelles de mesure vues précédemment comme le passage d'un taux de scolarisation à un revenu scolaire moyen.

2.1 ÉCHELLE DE MESURE ET REVENU SCOLAIRE MOYEN

Nous commençons par nous intéresser à une seule cohorte. Nous essayons donc de comparer le revenu scolaire moyen des différentes C.S.P. pour la cohorte au niveau N au temps T_1. L'étude de l'évolution entre T_1 et T_2 sera abordée au paragraphe suivant lorsque nous nous demanderons comment traduire qu'une évolution se fait à inégalités constantes.

Nous avons rapproché la variable latente «□revenu scolaire□» de la variable donnant le plus haut niveau éducatif atteint. Cette variable discrète est souvent quantifiée par les sociologues à partir d'un ordre voire d'un nombre d'années d'étude. Nous allons travailler sur des «□revenus scolaires□» continus. Il est toujours possible de considérer que les diplômes, une fois ordonnés, correspondent à des intervalles. Ce passage du discret au continu simplifie l'exposé sans nuire à l'étude des méthodes d'analyse de l'évolution des taux de scolarisation. Ajoutons que la hiérarchie des diplômes est souvent lue comme renvoyant à une hiérarchisation des emplois et, dans cette dernière, les revenus d'embauche qu'ils procurent ne constituent pas une dimension négligeable.

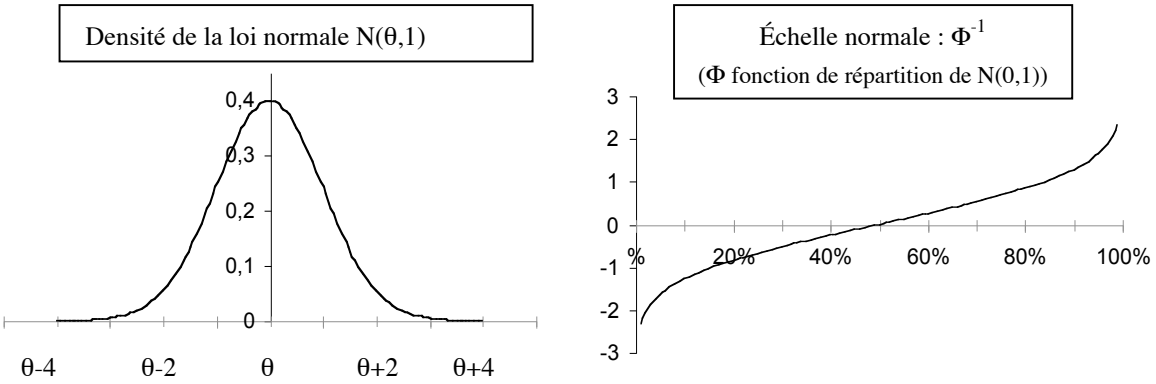
---

[18] La répartition des diplômes dans les différentes C.S.P. peut nous aider dans ce choix.

[19] Le paramètre θ indexe une distribution définie par sa fonction de répartition $F(\theta,t)$. C'est la probabilité que la variable latente prenne une valeur inférieure à t lorsqu'elle suit la distribution correspondant à θ. La décroissance stricte suppose que pour tout t et pour tout $\theta < \theta'$: $F(\theta,t) > F(\theta',t)$.

La première hypothèse que nous allons faire sur la distribution de la variable latente est celle de la normalité des revenus scolaires dans chaque C.S.P.. C'est un exemple sûrement peu réaliste mais il renvoie à des calculs bien connus. Pour la C.S.P. notée C_i on suppose que le revenu scolaire se distribue suivant une loi normale de moyenne θ_i et de variance σ_1_. Le paramètre θ_i est l'indice qui nous intéresse, la moyenne du revenu scolaire dans C_i. La variance σ_1_ est le deuxième paramètre inconnu indispensable pour définir une loi normale. Ne possédant que le taux de scolarisation pour un seul niveau nous sommes obligés de supposer la variance identique dans chaque C.S.P. Cette hypothèse n'est plus nécessaire lorsqu'on prend en compte les taux de scolarisation pour deux niveaux. Notons $R_1(N)$ le revenu scolaire correspondant au niveau N au temps T_1 et p_i le taux de scolarisation à ce niveau N dans la catégorie C_i. p_i est la fréquence des individus de C_i ayant un revenu scolaire supérieur à $R_1(N)$. Si Φ désigne la fonction de répartition de la loi normale centrée réduite, on a la propriété $\Phi[(R_1(N) - \theta_i)/\sigma_1] = 1 - p_i$ donc $\theta_i = \sigma_1\ \Phi^{-1}(p_i) + R_1(N)$. La moyenne θ_i n'est définie qu'à un changement d'origine et d'unité près puisqu'il en est ainsi de la variable latente.

La fonction Φ^{-1} étant continue et strictement croissante elle définit une échelle de mesure non bornée. Il en est de même de $\theta_i = f(p_i) = \sigma_1\ \Phi^{-1}(p_i) + R_1(N)$. Ce qui nous intéresse ce sont les écarts entre les θ_i, ils ne sont pas influencés par $R_1(N)$. Nous prendrons la même convention que précédemment : f(50 %) = 0, ce qui impose $R_1(N) = 0$ car $\Phi^{-1}(50\ \%) = 0$. Nous choisissons une unité de telle sorte que la variance σ_1_ des revenus, commune à toutes les C.S.P., soit égale à 1. Dans ces conditions, s'intéresser au revenu moyen c'est prendre pour échelle de mesure $f = \Phi^{-1}$. Le Graphique 4 nous donne à voir cette échelle ainsi que la densité de la loi normale de moyenne θ et de variance 1.

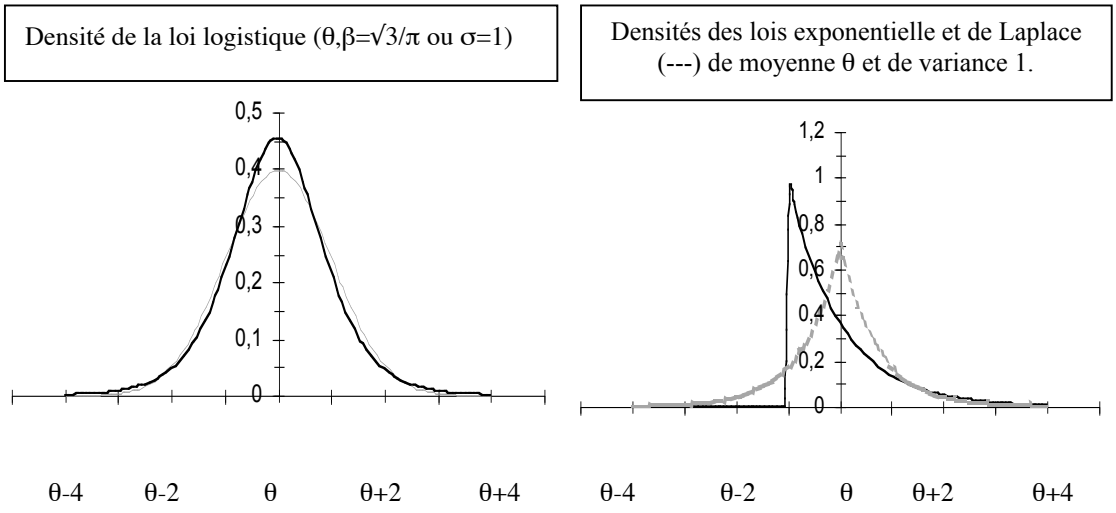


Graphique 4

Essayons maintenant de retrouver les échelles vues en 1, à partir de cette perspective. Il faut trouver la bonne famille de lois pour la variable latente.

L'échelle du taux logistique, $f(p) = \ln(p/(1 - p))$, s'obtient en considérant une variable latente suivant sur chaque C.S.P. une loi logistique de paramètres (θ,β)[20], elle est de moyenne θ et d'écart type $\sigma = \beta\pi/\sqrt{3}$. Comme le montre la figure ci-dessous la densité de la loi logistique de moyenne θ et d'écart type $\sigma_1 = 1$, c'est-à-dire $\beta = \sqrt{3}/\pi$, est proche de la loi normale de moyenne θ et d'écart type $\sigma_1 = 1$ (cette dernière est en trait clair).

Si θ_i désigne la moyenne du revenu scolaire dans la catégorie C_i, sous l'hypothèse d'une distribution logistique, on obtient comme pour la loi normale une échelle de mesure[21] $\theta_i = f(p_i) = \beta\ F_L^{-1}(p_i) + R_1(N)$ qui transforme le taux de scolarisation p_i en θ_i ($F_L(t) = 1/[1 + \exp(-t)]$ est la fonction de répartition de la loi logistique centrée $\theta = 0$ et de paramètre $\beta = 1$). Pour que cette échelle vérifie la convention $f(50\ \%) = 0$ il faut prendre $R_1(N) = 0$ car $F_L(0) = 50\ \%$. Avec $\beta = 1$ on obtient $f(p) = F_L^{-1}(p) = \ln(p/(1 - p))$[22] qui est le taux logistique. Pour comparer avec l'échelle normale il est préférable d'avoir aussi une variable latente de variance 1, ce qui s'obtient en prenant $\beta = \sqrt{3}/\pi$. L'échelle correspondante $(\sqrt{3}/\pi)\ln(p/(1 - p))$ est équivalente à l'échelle logistique, elle est très proche de l'échelle normale du Graphique 4 (nous ne l'avons pas tracée car la précision de la figure donne des courbes qui se recouvrent presque). Nous retrouvons ici une propriété bien connue en régression logistique : la proximité des modèles Probit et Logit (cf. par exemple [D.R. Cox, p. 29], [X. Guyon, 2001, p. 95]).



Graphique 5
(distributions conduisant aux échelles logistiques [ln(p/(1 - p)], de rapport [ln(p)] et des deux rapports [ln(p/0.5) si $0 < p \leq 0.5$ et $-\ln((1 - p)/0.5)$ si $0.5 \leq p < 1$])

[20] Si la variable X suit une loi logistique de paramètres ($\theta = 0$, $\beta = 1$), sa densité est égale à $\exp(x)/[1 + \exp(x)]^2$. La variable $Y = \beta X + \theta$ suit la loi logistique de paramètres (θ,β) avec $\beta > 0$.

[21] $F_L(t) = 1/[1 + \exp(-t)]$ est la fonction de répartition de la variable X de la note précédente. Dire que $Y = \beta X + \theta_i$ a une probabilité p_i d'être supérieure à $R_1(N)$ c'est dire que X a une probabilité p_i d'être supérieure à $(R_1(N) - \theta_i)/\beta$. On a donc $F_L((R_1(N) - \theta)/\beta) = 1\text{-}p_i$ ou $F_L(-(R_1(N) - \theta_i)/\beta) = p_i$ (la densité de X est symétrique) et $\theta_i = \beta\ F_L^{-1}(p_i) + R_1(N)$.

[22] $1/[1 + \exp(-t)] = p \Leftrightarrow \exp(-t) = (1/p) - 1 \Leftrightarrow t = \ln(p/(1-p))$.

L'échelle des différences consiste simplement à conserver les taux de scolarisation. Elle correspond à une variable latente de loi uniforme sur l'intervalle de centre θ et de longueur ℓ, c'est-à-dire de moyenne θ et d'écart-type $\ell/2\sqrt{3}$. Dans ce modèle la moyenne θ_i du revenu scolaire dans la catégorie C_i et le taux de scolarisation p_i vérifient $\theta_i = f(p_i) = \ell\ F_U^{-1}(p_i) + R_1(N)$ (F_U est la fonction de répartition de la loi uniforme centrée $\theta = 0$ et de paramètre $\ell = 1$)[23]. On obtient une échelle bornée, pour retrouver exactement celle des différences, avec les conventions utilisées, il faut prendre $\ell = 100$ et $R_1(N) = 50$. Les lois uniformes ont peu de chances d'être un bon modèle pour la distribution d'un revenu scolaire. Le taux de scolarisation ne peut donc pas être considéré comme un revenu moyen.

Les échelles de rapport (ln(p) et -ln(1-p)) correspondent, elles, à des lois exponentielles[24]. La densité exponentielle de moyenne θ et de variance 1, représentée dans le Graphique 5, conduit à l'échelle de rapport ln(p). La deuxième densité représentée est celle d'une loi de Laplace ou doublement exponentielle de moyenne θ et de variance 1, elle donne l'échelle des deux rapports du Graphique 2 (p. 9)[25]. Ces trois modèles de distribution d'un revenu scolaire ne semblent pas très réalistes, les échelles correspondantes ne peuvent donc pas être interprétées comme des revenus moyens. De ce point de vue elles ne semblent pas très adaptées comme indice d'inégalités.

Le principe de dualité ne disqualifie pas, dans cette présentation, les deux échelles de rapport. En effet le taux de non scolarisation $q_i = 1-p_i$ dans la catégorie C_i est la fréquence des individus de C_i ayant un revenu scolaire inférieur à $R_1(N)$. C'est donc la même équation qui donne θ_i comme fonction de p_i ou $1-q_i$. La recherche du θ_i définissant la distribution de la variable latente aboutit au même résultat, qu'on le fasse à partir du taux de scolarisation au niveau N ou à partir du taux de non scolarisation à ce même niveau. L'échelle de mesure n'est pas simplement vue comme une fonction du taux de scolarisation, mais comme une estimation à partir de ce taux d'un indice de scolarisation plus adapté. Si on part du taux de non scolarisation c'est ce même indice que l'on veut estimer, seul le processus d'estimation va changer. Dans le principe de dualité énoncé en 1, l'échelle de mesure est considérée comme valable pour les deux types de taux, de scolarisation et de non scolarisation. Quand on l'applique à ces derniers taux, ce n'est que dans l'interprétation que l'on fait intervenir la propriété de non scolarisation. Ceci conduit à un principe de dualité beaucoup plus fort, qui implique

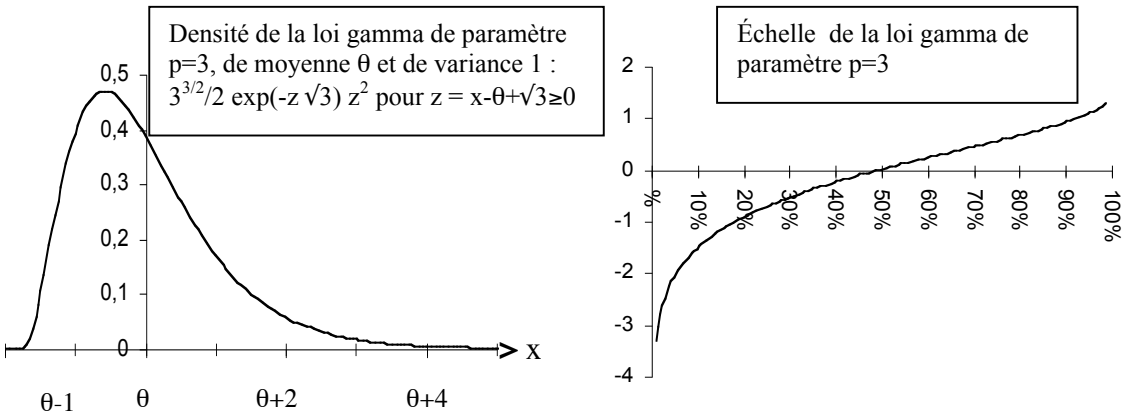
---

[23] Pour $-0.5 \le t \le +0.5$, $F_U(t) = t + 0.5$. Si X est une variable suivant cette loi, la variable $Y = \ell X + \theta$ suit la loi uniforme de paramètres (θ,ℓ) avec $\ell > 0$.

[24] Considérons la famille des lois engendrées par $Y = \sigma X + \theta$ avec $\sigma > 0$ et X de loi exponentielle centrée réduite, c'est-à-dire de densité exp(-x - 1) pour $x > -1$ (0 ailleurs). La fonction de répartition $F_X(t)$ de X est nulle pour $t < -1$ et égale à [1-exp(-t - 1)] sinon. Si θ_i est la moyenne du revenu scolaire et p_i le taux de scolarisation dans la catégorie C_i, avec la convention $\sigma = 1$ on a encore $F_X(R_1(N) - \theta_i) = 1-p_i$. On en déduit $\theta_i = \ln(p_i) + R_1(N) + 1$, pour avoir exactement l'échelle de rapport ln(p) il suffit de prendre $R_1(N) = -1$. L'autre échelle de rapport (celle des taux de non scolarisation) -ln(1 - p) s'obtient de façon semblable en partant d'une variable X de densité exp(x - 1) pour $x < 1$ (0 ailleurs).

[25] La démarche est semblable à celle de la note précédente, X suit une loi de Laplace centrée réduite de densité $(1/\sqrt{2})\exp(-|x|\sqrt{2})$ et de fonction de répartition : $F_X(t) = 0.5\exp(t\sqrt{2})$ si $t \le 0$, $F_X(t) = 1 - 0.5\exp(-t\sqrt{2})$ si $t \ge 0$. En prenant $R_1(N) = 0$ et $\sigma = \sqrt{2}$ l'équation classique $F_X((R_1(N) - \theta_i)/\sigma) = 1 - p_i$ se traduit par $\theta_i = \ln(p_i) - \ln(50\ \%)$ si $p_i \le 50\ \%$ et $\theta_i = \ln(50\ \%) - \ln(1 - p_i)$ si $p_i \ge 50\ \%$. C'est l'échelle des 2 rapports, en multipliant par 2 on obtient le Graphique 2.

(cf. 1) une symétrie de l'échelle de mesure *f* par rapport au point (50 %, *f*(50 %)). Une échelle de mesure construite pour donner la moyenne du revenu scolaire aura cette propriété de symétrie si la densité de la variable latente est symétrique[26]. C'est le cas pour l'échelle normale, l'échelle logistique et celle des deux rapports. Cette propriété n'est plus indispensable avec notre point de vue. On peut faire intervenir des distributions non symétriques. En particulier des distributions plus étalées du côté des revenus scolaires élevés que des bas revenus. Cette forme de distribution est suggérée par la répartition d'une classe d'âge suivant le diplôme obtenu, dans différentes C.S.P. (cf. par exemple [C. Thelot, L.-A. Vallet, 2000, p. 11]). Comme exemple nous pouvons considérer les distributions gamma[27]. Le Graphique 6 donne à voir la forme de la densité d'une loi gamma de paramètre p = 3, de moyenne θ et de variance 1. Si on impose cette forme de distribution à la variable latente revenu scolaire sur les individus de C_i, on obtient l'échelle de mesure[28] tracée dans le Graphique 6.



Graphique 6

Cette échelle se comporte un peu comme l'échelle des rapports ln(p) en donnant plus d'importance à un même écart des taux de scolarisation quand il est obtenu avec des petits taux. Ceci provient de la forme de la densité qui a très vite tendance à diminuer lorsque le revenu scolaire augmente. La fréquence des différents revenus scolaires est une fonction croissante sur les revenus les plus bas, mais elle devient vite décroissante pour les autres revenus. Le cas extrême est celui de l'échelle des rapports ln(p), elle correspond à la densité exponentielle qui n'a même pas de phase de

[26] La densité φ(x) vérifie φ(x) = φ(y) si le centre (x + y)/2 est la moyenne θ.

[27] La loi gamma de paramètres p > 0 et β > 0 a une densité non nulle pour x > 0, égale à : $[1/(\beta^p \Gamma(p))] \exp(-x/\beta) x^{p-1}$. Sa moyenne et sa variance sont respectivement égales à pβ et $p\beta^2$ (cf. [P. Tassi, 1989, p. 17]).

[28] La variable latente revenu scolaire est supposée avoir sur C_i la même loi que $X + \theta_i$, avec X de loi gamma centrée réduite de paramètre p = 3 et β = 1/√3 car on suppose la variance σ_1^2 égale à 1). Si F_X est la fonction de répartition de X on a comme précédemment $F_X(R_1(N) - \theta_i) = 1 - p_i$ donc $\theta_i = f(p_i) = R_1(N) - F_X^{-1}(1 - p_i)$ (la densité de X n'étant pas symétrique on n'a pas $F_X^{-1}(p_i) = - F_X^{-1}(1 - p_i)$). L'échelle f n'étant pas bornée nous prenons la convention f(50 %) = 0, c'est-à-dire $R_1(N) = F_X^{-1}(50\ \%) \cong -0.18818$. X ayant une loi gamma de paramètres p = 3 et β = 1/√3, translatée de -√3, on peut se servir pour les calculs de la fonction de répartition G de cette loi gamma. G^{-1} est donné par beaucoup de logiciels statistiques. On a $\theta_i = f(p_i) = G^{-1}(0.5) - G^{-1}(1 - p_i)$.

croissance (cf. le Graphique 5, p. 18). La loi exponentielle de variance 1 est d'ailleurs un cas particulier de loi gamma[29].

Les lois gamma ne sont pas les seules lois dissymétriques pouvant être utilisées. Nous les avons choisies parce qu'elles contiennent comme cas particulier la loi exponentielle. Dans le cas des variables intrinsèquement positives on utilise souvent les lois de Pareto-Levy ou les lois log-normales. Nous ne multiplierons pas ici les exemples, pour choisir les distributions de la variable latente il faudrait faire une étude des données sur les niveaux de sortie du système scolaire dans différentes catégories socio-professionnelles.

2.2. TRADUIRE LA CONSTANCE DES INÉGALITÉS

Nous introduisons maintenant la cohorte qui atteint le niveau scolaire N au temps T_2 afin d'étudier l'évolution du revenu scolaire entre T_1 et T_2. Au niveau scolaire N on dispose pour chacune des cohortes des taux de scolarisation dans les différentes C.S.P. retenues. Pour la C.S.P. C_i ils sont notés p_{i1} et p_{i2}. Le revenu scolaire correspondant au niveau N a pu évoluer entre T_1 et T_2, ce qui nous oblige à deux notations $R_1(N)$ et $R_2(N)$.

Notre étude des inégalités scolaires est basée sur les estimations des différents revenus scolaires moyens θ_{ij}, indexés par la C.S.P. C_i et la cohorte correspondant à T_j. Cette estimation dépend de la forme de la distribution du revenu scolaire que nous postulons sur les sous-populations étudiées. Nous en avons vu différents exemples au paragraphe précédent. Ils postulaient tous que ces distributions s'obtenaient par changement d'origine et d'unité à partir d'une distribution centrée réduite particulière (normale, logistique, exponentielle, gamma, …). Dans la suite nous définirons cette distribution choisie par sa fonction de répartition, notée F. Si σ_j^2 est la variance commune à toutes les C.S.P. au temps T_j, nous avons vu que l'estimation de θ_{ij} est donnée par $\theta_{ij} = R_j(N) - \sigma_j F^{-1}(1 - p_{ij})$[30].

Nous devons maintenant choisir les unités (σ_1 et σ_2) et les origines ($R_1(N)$ et $R_2(N)$) afin que les revenus scolaires soient comparables en T_1 et T_2. Aux temps T_j les inégalités entre les C.S.P. choisies s'expriment dans la variabilité des moyennes θ_{ij}, c'est une variabilité inter. Pour juger de la force des différences ($\theta_{i'j} - \theta_{ij}$) qui construisent cette variabilité, on prend souvent comme point de comparaison la variabilité des revenus à l'intérieur des C.S.P., elle est ici caractérisée par σ_j. Pour rendre comparable la variabilité des θ_{i1} à celle des θ_{i2} il apparaît donc nécessaire d'imposer $\sigma_1 = \sigma_2$, que nous prendrons égaux à 1 comme précédemment. La différence entre deux catégories ($\theta_{i'j} - \theta_{ij}$) au temps T_j ne dépend pas de $R_j(N)$, ce qui nous a permis au paragraphe précédent d'imposer à $R_j(N)$ une convention quelconque. Les évolutions des moyennes ($\theta_{i2} - \theta_{i1}$) dépendent, elles, de ($R_2(N) - R_1(N)$). Elles sont toutes translatées de la même quantité, qui n'intervient plus dans ce qui nous intéresse fondamentalement, les

[29] Il faut prendre $p = 1$ et $\beta = 1$ (cf. [P. Tassi, 1989, p. 17]).

[30] C'est une conséquence de l'équation $F((R_j(N) - \theta_{ij})/\sigma_j) = 1 - p_{ij}$, F étant continue et strictement croissante.

différences d'évolution. Afin de faciliter la lecture des différences intercatégories et des évolutions, il est pertinent d'imposer une moyenne nulle aux θ_{ij} d'une même cohorte : $\theta_{.1} = \theta_{.2} = 0$[31].

Avec ces conventions les inégalités restent constantes entre T_1 et T_2 si dans chaque C.S.P. les moyennes des revenus ne changent pas : $\theta_{i1} = \theta_{i2}$. C'est-à-dire qu'il existe pour les revenus scolaires au temps T_2 une unité et une origine qui rendent la représentation simultanée des distributions de ces revenus dans les différentes C.S.P. identique à celle au temps T_1. Les positions relatives des individus de ces différentes C.S.P. restent inchangées. Cette traduction de la constance des inégalités est équivalente[32] à celle correspondant à l'additivité (cf. 1.) des mesures obtenues avec une échelle liée à la fonction de répartition centrée réduite F, choisie pour modéliser la variable latente revenu scolaire.

Dans notre exemple de départ il n'y a que deux C.S.P. et deux cohortes, les conventions choisies entraînent $\theta_{11} = -\theta_{21}$ et $\theta_{12} = -\theta_{22}$. L'inégalité, au temps T_1, entre les fils de cadres et les fils d'ouvriers s'exprime dans la différence des revenus scolaires moyens de ces deux catégories : $\theta_{11} - \theta_{21} = 2\,\theta_{11}$. Au temps T_2 cette mesure de l'inégalité prend la valeur : $\theta_{12} - \theta_{22} = 2\,\theta_{12}$. Le jugement sur l'évolution de l'inégalité dépend de la comparaison entre ces deux différences de moyennes. Le Tableau ci-dessous donne à voir cette évolution pour différents modèles, nous avons choisi trois formes de densités (normale, logistique, gamma) pour la distribution du revenu scolaire[33].

Distribution choisie pour le revenu scolaire	Différence entre les revenus scolaires moyens des fils de cadres et des fils d'ouvriers	
	Avant 1910	Entre 1935 et 1940
Loi normale	1.99	1.59
Loi logistique	2.24	1.48
Loi gamma (p = 3)	2.98	1.80

Ces trois modélisations conduisent à une évolution allant dans le même sens, l'écart entre fils de cadres et fils d'ouvriers diminue. On peut obtenir une évolution dans l'autre sens en choisissant un autre type de distribution des revenus, par exemple la distribution uniforme qui correspond à l'échelle des différences, mais ce modèle de répartition des revenus scolaires n'est pas très réaliste. Un autre exemple est celui de la

[31] Si le nombre des C.S.P. étudiées est I, on a : $\theta_{.j} = \Sigma_i\, \theta_{ij}/I$. Nous accordons à chaque classe le même poids pour éliminer les effets de l'évolution de la distribution des C.S.P. Dans le cas de la recherche d'un indice global d'inégalité on pourrait prendre en compte la répartition des C.S.P.

[32] Nous avons vu au paragraphe précédent que l'échelle de mesure associée à F est de la forme $f(p_{ij}) = c - F^{-1}(1 - p_{ij}) = c - z_{ij}$, c étant une constante donnant une propriété intéressante à l'échelle. On en déduit $\theta_{ij} = z_{.j} - z_{ij}$, ce qui correspond à la partie $(\alpha_i + \gamma_{ij})$ dans la décomposition classique $f(p_{ij}) = \mu + \alpha_i + \beta_j + \gamma_{ij}$. On a $\theta_{i1} = \theta_{i2}$ pour tout i si et seulement si $\gamma_{i1} = \gamma_{i2} = 0$ car $\gamma_{i1} + \gamma_{i2} = 0$ par convention.

[33] F étant la fonction de répartition choisie pour modéliser la distribution des revenus, posons : $z_{ij} = F^{-1}(1 - p_{ij})$. On en déduit $\theta_{11} - \theta_{21} = z_{21} - z_{11}$ et $\theta_{12} - \theta_{22} = z_{22} - z_{12}$ (voir la note précédente).

distribution d'une variable X telle que son opposée -X soit de loi gamma de paramètre p = 3 (la forme de sa densité s'obtient en prenant la densité symétrique par rapport à l'axe des ordonnées de celle du Graphique 6, p. 20). Cette densité s'étale uniquement sur les bas revenus, ce qui ne nous semble pas être un bon modèle pour les revenus scolaires (ceux ci ne sont pas condensés sur les hauts revenus).

Il nous faut maintenant juger de la force de cette évolution. La variable latente revenu scolaire étant exprimée dans des unités qui ne nous parlent pas, il est difficile de dire si le passage d'une différence de 1.99 à 1.59 (pour le modèle normal) est une diminution notable ou négligeable. Pour avoir un point de comparaison on peut partir du fait que la variance intraclasses (moyenne des variances à l'intérieur des deux C.S.P.) est par convention égale à 1 pour chacune des deux cohortes. Ceci nous permet de juger de l'évolution de la variance interclasses qui ne dépend que de la différence entre les revenus scolaires moyens des fils de cadres et des fils d'ouvriers[34]. Classiquement on regarde l'évolution de la part de variance expliquée[35] par les C.S.P.. Dans notre exemple la part de variance expliquée sous le modèle normal passe de 50 % à 39 %, cette diminution de 11 % exprime la perte d'influence des deux C.S.P. sur le revenu scolaire (avec les modèles logistique et gamma on obtient respectivement une diminution de 20 % et 24 %). C'est une manière de regarder l'évolution du lien existant entre la variable C.S.P. et la variable revenu scolaire, la part de variance expliquée est interprétée ici comme indice d'inégalité. Il a l'avantage de pouvoir se calculer sur plus de deux C.S.P.[36] et de donner ainsi un indice synthétique d'inégalité. Bien entendu c'est un indice d'inégalité parmi d'autres. Nous n'allons pas en proposer d'autres car nous avons déjà vu dans la discussion sur le débat passé qu'il est difficile d'en choisir un qui serait meilleur que tous les autres. D'autre part l'évolution de ces indices donne bien un sens global à l'évolution des inégalités, mais une interprétation fine de ce qui a changé nécessite des informations supplémentaires. Une information facilement interprétable sur l'évolution de l'inégalité entre les différents couples de C.S.P. est souvent très utile.

Si nous prenons au hasard un individu dans chacune des deux C.S.P. étudiées, la probabilité que celui de la première catégorie ait un revenu scolaire supérieur à l'autre fait partie de ces informations qui font sens. On peut même supposer qu'un « acteur rationnel » prendrait en compte cette information pour juger de l'inégalité entre ces deux catégories. Dominique Goux et Éric Maurin [1995, p. 85] parlent de « retracer l'expérience que font les individus de la hiérarchie scolaire, suivant le milieu dont ils sont issus », lorsqu'ils calculent les probabilités qu'un jeune issu d'une famille de cadre ou de chef d'entreprise ait un diplôme supérieur, égal ou inférieur, à un jeune issu d'un milieu modeste. Pour construire ces probabilités ils utilisent les enquêtes sur la formation et la qualification professionnelle (FQP) de l'INSEE, ils doivent aussi ordonner les différents diplômes. Cette variable diplôme, discrète et ordinale, est la variable réelle qui se rapproche le plus de notre variable latente continue « revenu

[34] Si la différence entre les deux moyennes est égale à d, la variance équipondérée de ces deux moyennes est $(d/2)^2$. On pourrait ne pas accorder le même poids aux deux C.S.P. en prenant le poids qu'elles ont dans chaque cohorte. En faisant cela on prendrait en compte le changement social dans la définition de l'inégalité au lieu d'essayer de le neutraliser.

[35] C'est le rapport entre la variance interclasses et la variance totale (variance interclasses + variance intraclasses).

[36] On a plusieurs choix possibles pour les poids accordés aux différentes C.S.P. : prendre les poids des différentes catégories dans chaque cohorte ou essayer de neutraliser l'effet de l'évolution de la structure sociale en associant à chaque C.S.P. un poids identique pour toutes les cohortes.

scolaire. D'une certaine façon nous essayons de retrouver ces probabilités sans avoir la richesse des informations apportées par une enquête FQP.

Le tableau ci-dessous donne les résultats obtenus sur l'exemple de départ pour différents modèles.

Distribution choisie pour le revenu scolaire	Probabilité pour qu'un fils de cadre ait un revenu scolaire supérieur à un fils d'ouvrier	
	Avant 1910	Entre 1935 et 1940
Loi normale	92/100	87/100
Loi logistique	95/100	86/100
Loi gamma (p=3)	98/100	91/100

Les trois modèles fournissent des résultats[37] assez proches, il y a une diminution de 5 à 9 pour cent mais l'inégalité des chances reste grande, un fils de cadre a encore presque 90 chances sur cent d'avoir un revenu scolaire supérieur à un fils d'ouvrier[38]. L'interprétation est plus pessimiste qu'avec l'évolution de la part de variance expliquée, calculée précédemment. Ceci nous montre encore une fois l'importance des indices choisis pour résumer une évolution.

3. CONCLUSION

Porter un jugement sur l'évolution des inégalités à partir de taux de scolarisation à un niveau donné présente plusieurs difficultés : le regard est local, il est focalisé sur un seul niveau ; avec l'évolution du système scolaire, ce niveau ne représente pas le même potentiel au cours du temps ; les mêmes différences de taux ne conduisent pas forcément au même jugement suivant que les taux utilisés sont petits ou grands. En faisant intervenir une variable latente, le « revenu scolaire », nous avons essayé de considérer une mesure de l'inégalité plus globale : les revenus scolaires moyens dans les différentes catégories sociales considérées. Le passage des taux de scolarisation aux moyennes peut aussi s'interpréter comme un changement d'échelle de mesure. De nombreuses échelles de mesure ont déjà été proposées afin de permettre la comparaison des évolutions de taux de scolarisation. Nous avons vu que seule l'échelle logistique peut être interprétée comme transformation donnant une moyenne de revenus. Elle permet donc deux interprétations, l'une locale et l'autre plus globale. La neutralisation de l'évolution du système scolaire se traduit comme dans le cas des échelles de mesure par l'additivité des moyennes. Dans ce cas les distributions des revenus scolaires de la deuxième cohorte sont celle de la première cohorte à un changement d'origine et d'unité près. Les individus des différentes catégories sociales se trouvent dans des positions

[37] F étant la fonction de répartition centrée réduite choisie pour modéliser la distribution des revenus, nous devons calculer la probabilité qu'une variable U = Y - X soit positive sous les conditions : X et Y sont indépendantes, X suit la loi F, Y a la loi de X + d où d est la différence entre les revenus scolaires moyens des deux C.S.P.. Si F_U désigne la fonction de répartition de la variable différence U, on doit calculer $P(U > 0) = 1 - F_U(0)$. Dans le cas du modèle normal le résultat se trouve facilement puisque U suit une loi normale de moyenne d et de variance 2. Dans le cas général $F_U(0)$ s'obtient en calculant l'intégrale de F(x - d) par rapport à la loi de X (cf. par exemple [A. Montfort, 1980, p. 100]).

[38] C'est l'ordre de grandeur trouvé par D. Goux, E. Maurin ([1995] et [1997]) pour la probabilité qu'un jeune adulte d'origine supérieure soit au moins autant diplômé qu'un jeune d'origine populaire, en 1970 et 1993.

relatives identiques. Cette neutralisation remplace la recherche, à des moments différents, de niveaux équivalents pour les potentialités qu'ils donnent à leur possesseur.

L'estimation des revenus scolaires moyens à partir des taux de scolarisation repose sur une modélisation de la variable latente «revenu scolaire». Nous avons dans cet article fait différentes hypothèses pour la forme de la distribution de cette variable sur les populations concernées. Dans chaque cas nous avons supposé que les variances des revenus scolaires des sous-populations appartenant à une même cohorte étaient identiques. C'est une hypothèse un peu restrictive, si on veut la supprimer et ainsi améliorer le modèle il faudrait avoir les taux de scolarisation pour un autre niveau d'étude. On pourrait alors estimer la moyenne et la variance des revenus scolaires sur chaque C.S.P. à partir des deux taux de scolarisation. L'évolution des inégalités s'exprimerait dans l'évolution des moyennes mais aussi dans celle des variances. Ce qui permettrait de rendre compte de l'évolution de la cohésion de chaque catégorie sociale face au système scolaire.

Le choix des distributions servant à modéliser la variable latente peut être guidé par les informations que nous donnent les enquêtes F.Q.P. sur la distribution des diplômes dans certaines cohortes. Ce choix peut ainsi en partie reposer sur des données réelles et non pas uniquement sur des propriétés abstraites jugées intéressantes comme dans le cas du choix d'une échelle de mesure. Ainsi l'échelle logistique possède de nombreuses propriétés mathématiques et des facilités d'interprétation à mettre à son actif, Henri Léridon et Laurent Toulemon [1997, p. 247] disent même que «toute échelle vérifiant quelques contraintes d'exhaustivité et de symétrie doit lui ressembler». Ceci explique le succès de la régression logistique et des modèles log-linéaires pour traiter de l'évolution d'un ensemble de pourcentages. Avec le point de vue que nous avons adopté l'échelle logistique (distribution logistique de la variable latente) est un modèle acceptable, elle a toutefois le défaut d'être symétrique ce qui ne semble pas très adapté au revenu scolaire. De plus les arguments en faveur de l'échelle logistique sont liés aux bonnes propriétés de la transformation qu'elle fait subir aux taux de scolarisation, or dans le débat qui nous occupe il ne faut pas oublier que le choix d'une échelle impose des propriétés à la notion de constance des inégalités traduite dans la propriété d'additivité. Suivant l'échelle retenue l'additivité donne des résultats différents sur ce que seraient les taux de scolarisation de la première cohorte, sous l'hypothèse d'une évolution du système scolaire à inégalité constante, si elle avait grandi en même temps que la deuxième cohorte. Dans notre exemple de départ, pour avoir des données additives, il faudrait que le taux de réussite des fils de «cadres» nés entre 1935 et 1940 ne soit pas de 62 % (les trois autres taux sont inchangés) mais de 86,6 % dans le cas logistique et de 99,9 % dans le cas de l'échelle gamma. Supposons que l'évolution de la scolarisation se traduise par une augmentation uniforme du niveau scolaire, une translation des revenus scolaires, on obtiendrait les distributions des revenus scolaires de la deuxième cohorte dans les différentes catégories en translatant d'une même quantité les distributions de la première cohorte. Si la forme de ces distributions n'est pas proche de la distribution qui définit l'échelle de mesure on trouvera une diminution ou une augmentation de l'inégalité. Les bonnes propriétés de l'échelle logistique dans l'analyse des proportions ne nous semblent pas clore le débat sur la modélisation de l'évolution comparée des inégalités des taux de scolarisation.

Nous avons essayé de montrer les difficultés d'analyse de l'évolution de l'inégalité des taux de scolarisation en utilisant une modélisation simple de la variable latente «revenu scolaire», pour ne pas cacher ces difficultés par la mise en place d'un modèle plus complexe. Ce modèle simplifié permet de construire des sociétés idéales typiques à

inégalités constantes, sur lesquelles on peut étudier les propriétés des outils d'analyse classiquement utilisés. Nous avons proposé quelques pistes pour complexifier ce modèle. Essayer d'appliquer ce point de vue à des exemples plus étoffés serait une première étape indispensable, mais il s'agirait d'un autre article qui nécessiterait une analyse fine des statistiques disponibles. Le choix des distributions de la variable latente « revenu scolaire » suggéré par ces analyses conduirait à réaliser un travail statistique plus ou moins important, suivant la complexité du modèle final.

BIBLIOGRAPHIE

ANTONIADIS A., BERRUYER J., CARMONA R., *Régression non linéaire et applications*, Paris, Economica, 1992.

BARBUT M., « Note sur quelques indicateurs globaux de l'inégalité : C. Gini, V. Pareto, P. Lévy », *Revue française de sociologie* 25 (4), 1984(a), p. 609-622.

BARBUT M., « Sur quelques propriétés élémentaires des fonctions de concentration de C. Gini », *Mathématiques et Sciences humaines,* n° 88, 1984(b), p. 5-19.

BARBUT M., « Sur les indicateurs de l'inégalité : croissance logistique et mesure de l'inégalité, et de quelques effets « paradoxaux » dans la comparaison des inégalités », *Mathématiques et Sciences humaines*, n° 90, 1985, p. 5-17.

BOUDON R., *L'inégalité des chances. La mobilité sociale dans les sociétés industrielles*, Paris, Armand Colin, 1973.

CHRISTENSEN R., *Log-linear models*, New York, Springer-Verlag, 1990.

COMBESSIE J.-C., « L'évolution comparée des inégalités : problèmes statistiques », *Revue française de sociologie* 25 (2), 1984, p. 233-254.

COMBESSIE J.-C., « Paradoxes des fonctions de concentration de C. Gini », *Revue française de sociologie* 26 (4), 1985, p. 653-658.

COX D. R., *The analysis of binary data*, London, Methuen, 1970.

EURIAT M., THELOT C., « Le recrutement social de l'élite scolaire en France. Évolution des inégalités de 1950 à 1990 », *Revue française de sociologie* 36 (4), 1995, p. 403-438.

FLORENS J.-P., « Inégalité et dépendance statistique », *Revue française de sociologie* 25 (2), 1984, p. 255-263.

GOUX D., MAURIN E., « Origine sociale et destinée scolaire. L'inégalité des chances devant l'enseignement à travers les enquêtes Formation-Qualification Professionnelle 1970, 1977, 1985 et 1993 », *Revue française de sociologie* 36 (1), 1995, p. 81-121.

GOUX D., MAURIN E., « Démocratisation de l'école et persistance des inégalités », *Économie et statistique*, 306, 1997, p. 27-39.

GREMY J.-P., « Sur les différences entre pourcentages et leur interprétation », *Revue française de sociologie*, 25 (3), 1984, p. 396-420.

GUYON X., *Statistique et économétrie : du modèle linéaire… aux modèles non-linéaires*, Paris, Ellipses, 2001.

LERIDON H., TOULEMON L., *Démographie : approche statistique et dynamique des populations*, Paris, Economica, 1997.

MERLLIE D., « Analyses de l'intéraction entre variables. Problème statistique ou sociologique ? », *Revue française de sociologie* 26 (4), 1985, p. 629-652.

MONFORT A., *Cours de probabilités*, Paris, Economica, 1980.

PREVOT J., « À propos d'indices et de comparaisons de proportions », *Revue française de sociologie* 26 (4), 1985, p. 601-628.

SUSSMAN L., *Objectifs sociaux et planification de l'enseignement*, Paris, OCDE, 1969.

TASSI P., *Méthodes statistiques*, Paris, Economica, 1989.

TEROUANNE E., « Distorsion entre deux distributions d'une variable nominale », *Mathématiques, Informatique et Sciences humaines,* n°131, 1995, p. 29-38.

THELOT C., VALLET L.-A., « La réduction des inégalités sociales devant l'école depuis le début du siècle », *Économie et statistique* 334, 2000, p. 3-32.

VALEYRE A., « Formes et propriétés des indices d'inégalité entre proportions », *Mathématiques, Informatique et Sciences humaines*, n° 132, 1995, p. 13-37.

VALLET L.-A., « L'évolution de l'inégalité des chances devant l'enseignement. Un point de vue de modélisation statistique », *Revue française de sociologie* 29 (3), 1988, p. 395-423.

VALLET L.-A., « Quarante années de mobilité sociale en France. L'évolution de la fluidité sociale à la lumière de modèles récents », *Revue française de sociologie* 40 (1), 1999, p. 5-64.

WESTERGAARD J., LITTLE A., "Educational opportunity and social selection in England and Wales. Trends and policy implications", *Social Objectives in Educational Planning*, Paris, OCDE, 1967.