

UN MODÈLE STATISTIQUE GLOBAL POUR LES CONSEILS D'ORIENTATION

Texte préparé pour le XIe Colloque de l'ADMÉE
"L'évaluation dans le processus d'orientation"
qui s'est tenu à l'Office de la Recherche Pédagogique
à Tramelan
22-23-24 septembre 1997

Jean CARDINET

UN MODÈLE STATISTIQUE GLOBAL POUR LES CONSEILS D'ORIENTATION

Texte préparé pour le XIe Colloque de l'ADMÉE
"L'évaluation dans le processus d'orientation"
qui s'est tenu à l'Office de la Recherche Pédagogique
à Tramelan
22-23-24 septembre 1997

Jean CARDINET

IRDP

Institut de recherche et de documentation pédagogique
Faubourg de l'Hôpital 43 - Case postale 54 - CH 2007 Neuchâtel 7
Tél: ++41 32 889 69 70 / Fax: ++41 32 889 69 71 / Internet: <http://www.unine.ch/irdp/>

CARDINET, Jean. - Un modèle statistique global pour les conseils d'orientation : texte préparé pour le XIe Colloque de l'ADMEE « L'évaluation dans le processus d'orientation » qui s'est tenu à l'Office de la recherche pédagogique à Tramelan, 22-24 septembre 1997 / Jean Cardinet. - Neuchâtel : Institut de recherche et de documentation pédagogique (IRDP), 1998. - 17 p. ; 30 cm. - (Recherches ; 98.101) (Cahier du GCR ; 34). - Bibliogr. p. 17
CHF 4.-

Modèle statistique
Orientation pédagogique
Orientation professionnelle

Méthode statistique
Prise de décision

Les cahiers du GCR/SSRE présentent d'une part, les contributions exposées lors des rencontres plénières du Groupe des Chercheurs Romands, d'autre part des réflexions, des articles et des comptes rendus de recherche rédigés par des chercheurs en sciences de l'éducation, extérieurs à l'IRDP.

Les réflexions développées dans les rapports par des personnes extérieures à l'IRDP n'engagent que la responsabilité de leur(s) auteur(s) et non celle de l'IRDP.

La reproduction totale ou partielle des publications de l'IRDP est en principe autorisée à condition que leur(s) auteur(s) en ai(en)t été informé(s) au préalable et que les références soient mentionnées.

UN MODÈLE STATISTIQUE GLOBAL POUR LES CONSEILS D'ORIENTATION

Jean Cardinet

Résumé

Conseiller une orientation, scolaire ou professionnelle, suppose que l'on a une base théorique pour déterminer le "bon" choix. Un modèle quantitatif est présenté, s'inspirant de la théorie de la décision. Il tient compte de l'instance de décision (fixant par là l'échelle de valeurs à considérer), puis du degré de connaissance que l'on a de la personne à orienter, et enfin de la possibilité d'estimer, ou non, son succès probable. Six cas différents en résultent, qui font chacun appel à une méthode statistique particulière.

A GLOBAL STATISTICAL MODEL FOR GUIDANCE

Jean Cardinet

Summary

Scholastic or vocational guidance implies that the counselor can determine the "good" choice, on some theoretical basis. A quantitative model is presented, based on decision theory. It takes into account who makes the choice (and hence, which utility scale is to be considered), what degree of knowledge the counselor has of his client, and finally whether it is possible or not to predict his probable success. Six different cases result from these various combinations, each one calling for a particular statistical method.

UN MODÈLE STATISTIQUE GLOBAL POUR LES CONSEILS D'ORIENTATION

Abrégé

Décider de l'orientation des élèves sur la base d'une succession d'examens de sélection représente une des plus mauvaises stratégies possibles. Une bonne utilisation des compétences implique plutôt d'exploiter au mieux les points forts de chaque personne, en renonçant à classer tous les candidats sur une échelle unique. Ceci n'est qu'un exemple de ce que peuvent démontrer les analyses quantitatives des procédures d'orientation.

De telles études supposent que l'on se donne un modèle statistique de la situation d'orientation, qui soit assez englobant pour s'appliquer à des situations très différentes et assez riche pour mettre en évidence l'effet de plusieurs paramètres. C'est pourquoi le modèle qui sera présenté fait varier l'instance de décision (et par là l'échelle de valeurs à prendre en compte), puis la connaissance que l'on a de l'étudiant, et enfin la possibilité d'estimer son succès probable.

L'instance de décision peut être la personne individuelle, qui va valoriser telle ou telle orientation, selon son échelle de valeurs propre, ou bien le système scolaire, dont il s'agit alors d'augmenter le rendement moyen, ou bien une combinaison des deux, où l'individu bénéficie en retour de son apport à la société.

La connaissance que l'on a de l'étudiant permet de l'attribuer à certaines catégories (psychologiques, sociales, etc.), soit avec certitude, soit en probabilité, soit ne le permet pas du tout, ce qui va jouer un rôle du fait que ces catégories interagissent avec les conditions d'enseignement, pour faciliter ou non l'apprentissage de l'étudiant et modifier sa réussite attendue.

Ce succès futur, enfin, représente l'élément essentiel de l'utilité du choix d'orientation effectué, (même si d'autres aspects, comme la satisfaction de l'intéressé, sont aussi à prendre en compte dans le résultat). Le modèle qui sera présenté envisage plusieurs possibilités concernant l'estimation de ce critère d'utilité attendue. Il peut admettre qu'il existe une relation multilinéaire (donnée par une équation de régression multiple) entre la série de prédicteurs, mesurés pour chaque étudiant, et le succès ou la satisfaction que ce dernier obtiendra dans chaque orientation possible. Le modèle peut admettre au contraire que le résultat attendu n'est connu que pour certaines catégories d'étudiants, ou seulement de façon globale pour tous ceux qui appartiennent à une catégorie donnée.

La combinaison de ces diverses possibilités conduit à présenter six procédures d'orientation différentes, affectées encore par la présence, ou non, de quotas de répartition à respecter.

L'important, pour les évaluateurs chargés de mettre en place des examens d'orientation, est de se rendre compte de cette diversité des procédures possibles et de choisir la logique qui correspond le mieux à leur situation (même s'il est clair qu'une quantification effective est impossible). En particulier, ils doivent éviter de s'en tenir à des "écrémages successifs" et valoriser en priorité les facteurs de réussite autres que le niveau scolaire général.

Les six modalités d'optimisation de l'orientation

<i>Succès estimable</i>	Catégorie d'appartenance		
	Certaine	En probabilité	Inconnue
<i>Pour tous</i>	Classement différentiel	Cl. combiné probabiliste	Classement minimax
<i>Dans la catégorie</i>	Classement modéré	Cl. combiné selon les risques connaissables	
<i>Globalement par catégorie</i>	Programmation linéaire	Fonctions discriminantes pondérées	

UN MODÈLE STATISTIQUE GLOBAL POUR LES CONSEILS D'ORIENTATION

Jean Cardinet
Neuchâtel

But de ce document

Qu'est-ce qu'une "bonne" orientation?

La raison d'être de ce texte, c'est le manque de cadre théorique précis pour guider le choix d'une orientation scolaire ou professionnelle. Comment reconnaît-on que tel choix est préférable à tel autre?

L'analyse effectuée par l'auteur, pendant l'année 1960, des conseils apportés par André Rey dans le cadre des consultations qu'il donnait devant ses étudiants à Genève n'a pas révélé de stratégie systématique. Il semble qu'il en était de même dans les Offices d'orientation de l'époque. Les conseillers essayaient de voir quels étaient les intérêts de leur client, lui suggéraient une profession en accord avec ses vœux et se contentaient de vérifier que le niveau scolaire et les aptitudes intellectuelles de l'intéressé permettaient d'espérer son succès dans la formation qu'il envisageait.

N'y aurait-il pas un avantage, cependant, à articuler de façon plus claire les degrés d'intérêt que peuvent susciter plusieurs choix possibles et les probabilités de réussite dans chacune de ces directions, pour comparer un peu plus précisément leurs utilités globales respectives?

Est-il correct aussi de faire comme si le marché du travail allait permettre de répondre positivement à toutes les demandes des jeunes? Ne faudrait-il pas aujourd'hui tenir compte des "quotas" existant de fait dans notre société? (Il faut plus de mécaniciens que d'étalagistes, par exemple.)

L'apport souhaité de la théorie de la décision

La référence à une "théorie des jeux" et notamment à sa forme simplifiée d'un jeu contre la Nature, tel que l'envisageait Laplace, ainsi qu'à la "théorie de la décision" qui en est un des chapitres, semble devoir faciliter l'intégration des diverses informations dont on dispose, relatives à l'individu et aux directions d'orientation envisagées, dans la ligne de l'ouvrage de Cronbach et Gleser (1965).

Selon ce modèle, il existe pour chaque choix possible une utilité attendue, moyenne des résultats que l'on obtiendrait en répétant le jeu un grand nombre de fois. On calcule cette utilité en explicitant les différentes conséquences qui pourraient s'en suivre si l'on faisait ce choix. Chaque conséquence a une utilité positive, ou un coût négatif, mais aussi une probabilité plus ou moins grande de se produire. On pondère (multiplie) alors les utilités prévues par la probabilité que cet événement arrive et on totalise les produits.

Il semble raisonnable de choisir l'option comportant la plus grande utilité attendue, même si le résultat effectif dans une circonstance particulière peut naturellement être très différent du résultat moyen calculé.

Dans certaines situations, il pourrait être plus prudent de chercher à éviter la perte maximale. (Voir plus bas le classement minimax.) Un tempérament joueur, au contraire,

pourrait vouloir tenter de décrocher le gros lot en choisissant l'option qui comporte la possibilité la plus avantageuse.

Pour le lecteur, c'est déjà un apport de la théorie de la décision d'attirer son attention sur ces trois principes de choix possibles: optimiste, pessimiste et pondéré. Pour un évaluateur appelé à conseiller de nombreux clients, ou à déterminer une stratégie valable pour toute une organisation, c'est naturellement l'utilité attendue en probabilité qui est à considérer.

Données supposées connues

Utilités subjectives et objectives

Le lecteur verra aussi dès le début ce qui limite l'application effective des formules envisagées: il faudrait pouvoir disposer des valeurs d'utilités dont on parle, ce qui n'est évidemment pas le cas. Pourtant nous sommes rarement dans l'ignorance totale. Telle option semble au sujet plus intéressante, ou plus prestigieuse, ou permettant une promotion plus rapide, ou connaissant moins le chômage, etc. que telle autre. On peut imaginer un instrument qui totaliserait ces valeurs d'utilité partielles et en tirerait une certaine quantification de l'utilité subjective.

On s'est souvent basé sur le salaire moyen que recevaient les membres d'une profession pour estimer l'utilité sociale objective d'un métier, mais il faut ensuite corriger cette estimation en faisant évaluer aussi la façon dont la personne considérée donne satisfaction dans son poste ou progresse dans sa formation. Le simple fait de se poser la question de ces utilités amène à analyser plus en profondeur les éléments à considérer dans ce qu'on appelle la réussite d'une orientation.

Si l'on voulait opérationnaliser vraiment le modèle, on serait amené à constituer des familles d'emplois et des voies générales de formation, regroupant les choix possibles en une dizaine de directions au plus, pour pouvoir maîtriser la complexité des problèmes d'estimation. Il faudrait alors pouvoir évaluer le niveau de compétence ou de réussite sur une échelle commune, pour les membres de chacune de ces orientations.

Prédicteurs

Les utilités dont on vient de parler sont celles dont font preuve les personnes déjà engagées avec succès dans une voie particulière. Dans le modèle psychométrique, elles sont données par les **critères** mesurant la réussite, en fin de formation ou dans la pratique ultérieure.

Comme ces valeurs ne sont pas connues au moment où la décision d'orientation doit être prise, on devra les estimer à partir des autres informations disponibles au départ sur les candidats: données biographiques, diplômes scolaires obtenus, tests d'aptitude et d'intérêt, questionnaires, etc. Ces variables constituent les **prédicteurs**.

Relations existant entre prédicteurs et utilités

D'innombrables recherches de psychologie différentielle appliquée ont établi les relations qu'on pouvait observer en général entre les prédicteurs habituels et la réussite dans chaque formation ou profession. Des équations de régression ont pu alors être calculées, permettant d'estimer à l'avance l'utilité probable de chaque candidat.

Ces relations ne pouvaient cependant être observées que pour des personnes déjà en place; d'où un problème au moment où un conseiller d'orientation veut les appliquer directement à une population tout venant.

Les recherches ont aussi établi que certaines catégories de personnes étaient plus prédictibles que d'autres, ou que les relations des prédicteurs à la réussite différaient selon l'origine scolaire ou sociale, etc. des candidats.

C'est pour ces deux raisons qu'on propose que la prédiction de réussite soit calculée en deux phases, la première étape consistant à déterminer, selon la catégorisation qu'on peut faire du sujet, quelle est l'équation de régression à utiliser dans la seconde étape.

Position du problème

Déterminer une "stratégie" qui maximise l'utilité attendue

On appelle stratégie une règle de décision, ou programme envisageant tous les cas possibles et disant quelle décision prendre dans chaque cas.

L'ambition du modèle présenté est d'esquisser le principe d'une démarche qui analyserait les informations recueillies sur chaque client, les traduirait en utilités attendues pour chaque orientation, puis en déduirait l'orientation à conseiller, compte tenu des caractéristiques du marché du travail.

Le conseil serait établi de façon à ce que la somme des utilités attendues, pour toutes les orientations indiquées, soit la plus élevée possible.

Choix de l'instance de décision

L'individu

Lorsqu'il est question d'orientation, on pense que c'est à la personne elle-même de choisir la voie de formation ou la profession qu'elle désire. C'est dire en même temps que ce sont les valeurs de l'individu qui doivent être prises en compte pour calculer les valeurs d'utilité de chaque orientation.

L'institution

Mais lorsqu'une sélection est nécessaire, y compris pour plusieurs voies à la fois, il faut un décideur institutionnel pour arbitrer la compétition. Ce sont alors les valeurs de l'institution et l'utilité pour elle, auxquelles on se réfère.

L'utilité combinée

Compte tenu de la compétition qui règne sur le marché de l'emploi, l'idée est défendue ici que les évaluateurs chargés de l'orientation devraient jouer le rôle d'arbitres et prendre en compte des valeurs d'utilité combinées, qui intègrent le bénéfice pour l'individu de l'amélioration qu'il peut apporter au fonctionnement économique et social (selon la logique des jeux coopératifs).

Choix de la catégorie d'appartenance

Exemples de catégories

Ghiselli (1960) a introduit dans la psychométrie un concept important, celui de "prédictabilité". Il a montré qu'il était possible de distinguer un groupe de personnes pour lesquelles la corrélation était élevée entre prédicteurs et critère, et un groupe "imprédictible". Une prédiction en deux temps de l'appartenance, puis du rendement, permet d'améliorer la sélection.

Les grandes universités américaines, qui doivent prédire la réussite de candidats venant de collèges différents, ont établi des équations de régression spécifiques pour chaque collège. Elles peuvent tenir compte ainsi des caractéristiques de l'enseignement donné et de la notation appliquée dans chacun d'eux.

Certaines études présentent la réussite scolaire des écolières comme mieux prédictible, ou comme résultant d'autres types de comportements, que celle des écoliers. La prédiction du succès serait donc améliorée si l'on calculait une régression séparément pour chaque groupe.

Les mécaniciens formés par un apprentissage en atelier et ceux qui sont formés en école n'ont pas développé les mêmes sortes de compétences. On peut faire l'hypothèse qu'ils aborderont différemment la même tâche et donc que les validités des tests seront différentes dans les deux groupes.

Un doctorat en médecine étant requis pour avoir le droit de pratiquer, ce diplôme ne peut pas différencier les médecins plus ou moins efficaces. Une prédiction de succès professionnel dans ce cas est forcément conditionnée en premier lieu par l'appartenance à la catégorie des médecins.

Pourquoi tenir compte de l'appartenance

Les exemples précédents montrent que la prédiction de l'utilité peut difficilement se passer souvent d'une prise en compte des groupes d'appartenance. Mais d'autres raisons encore militent pour l'introduction de cette préoccupation en orientation.

La démarche intuitive des conseillers d'orientation qui reconnaissent des profils de résultats et d'intérêts "typiques" d'un mécanicien-auto, par exemple, s'apparente en fait à la prédiction d'une catégorie professionnelle, plus qu'à celle du succès dans cette profession. Le modèle recherché doit le permettre aussi.

La matrice de variances-covariances totale des prédicteurs et des critères peut être scindée en deux parties, intergroupes et intragroupes. Les études de validité et les équations de régression habituelles reposent sur l'analyse des différences intragroupes, alors que les fonctions discriminantes exploitent les différences intergroupes. Il est souhaitable que le modèle statistique envisagé tienne compte de ces deux types d'information à la fois.

Catégories connues, probables, ou inconnues

Pour couvrir l'ensemble des cas possibles, le modèle distinguera trois situations:

- les catégories sont connues de façon sûre au moment de l'évaluation des résultats. C'est possible s'il s'agit de données biographiques, ou de diplômes sanctionnant des études antérieures, par exemple.
- les catégories ne sont pas connues de façon assurée, mais il est possible d'en estimer la probabilité a posteriori. On dira ainsi que sur 100 personnes ayant ce

profil de résultats, 50% deviennent des universitaires, 30% des cols blancs et 20 % des cols bleus.

- les catégories ne sont pas connaissables au moment de l'examen, qu'elles existent effectivement ou non. Les futurs cas d'abandon, par exemple, sont généralement difficiles à déceler au moment où des candidats se présentent, parce que les raisons d'abandon peuvent être multiples.

Possibilités d'estimation de l'utilité

Estimation individuelle possible hors catégories

La psychométrie classique suppose que les relations entre prédicteurs et critères, trouvées à l'intérieur d'une profession, sont valables de façon générale et peuvent être appliquées à n'importe quelle personne, même non encore formée dans cette profession. Une équation de régression multiple permet d'estimer à l'avance son efficacité probable dans cette direction. C'est que la psychotechnique prend comme prédicteurs des tests d'aptitudes, généralement peu influençables par une formation professionnelle particulière, et qu'elle vise surtout des professions où la formation est rapide.

Estimation individuelle possible par catégorie

En réalité, pour les postes de cadres en particulier, mais aussi pour beaucoup de professions de haut niveau, les critères de réussite ne sont obtenables que pour des personnes très sévèrement sélectionnées au préalable. L'équation qui permet de prédire ces critères à partir de certains prédicteurs n'est logiquement pas applicable hors de cette sous-population spécifique. L'utilité à attendre d'un certain choix est alors conditionnée par l'appartenance à la catégorie professionnelle en question, ce qui justifie une prédiction en deux temps, de l'appartenance d'abord, de la réussite ensuite.

Estimation globale possible par catégorie

Il arrive même que la dispersion des aptitudes et des intérêts soit si faible à l'intérieur de certaines orientations que l'homogénéité interne du groupe empêche de distinguer quels membres manifestent plus d'utilité que les autres. La démarche d'orientation ne peut faire mieux alors que d'attribuer une utilité globale identique à tous ceux qui ont choisi cette orientation.

Le plus souvent, c'est simplement l'absence de données sur les utilités individuelles au sein de la profession ou de la voie de formation choisie, qui amène à se contenter d'une estimation globale.

Calcul des utilités attendues de chaque orientation

Solution générale

Le principe de la stratégie recherchée est d'estimer quelle est la valeur attendue de chaque possibilité envisagée, en tenant compte de toutes les conséquences envisageables d'un tel choix. On effectuera ensuite la répartition entre les orientations de manière à maximiser la somme de toutes les valeurs attendues d'utilité pour les options choisies.

L'équation générale a déjà été esquissée au début de ce texte. Il s'agit de sommer les utilités pour chaque état de nature ou événement envisageable, chacun d'eux étant multiplié par la probabilité a posteriori qu'il se produise.

Par exemple, il se peut qu'un étudiant sur quatre ne termine pas sa formation dans la voie qu'il avait commencée, sans qu'on sache prédire cet abandon, dû à une multiplicité de causes aléatoires. La probabilité d'abandon est donc $p_a = 0.25$, la probabilité complémentaire de finir ses études étant $p_f = 0.75$. Si l'on chiffre le coût de l'abandon à c_a (valeur négative) et l'utilité de la formation terminée à u_f pour le candidat envisagé, l'utilité moyenne attendue si l'on choisit cette orientation est:

$$\text{Utilité moyenne attendue} = p_a \cdot c_a + p_f \cdot u_f$$

C'est cette valeur qu'il faudra comparer aux utilités moyennes attendues des autres choix.

Il y aurait, bien sûr, de nombreuses façons de découper les catégories de candidats et les éventualités à considérer. En principe, il faudrait utiliser les regroupements qui donnent le mieux prise sur le réel, en permettant de prévoir les différences d'utilité les plus grandes. Mais comme on le verra en conclusion, le modèle n'a pas besoin d'être vraiment concrétisé pour être déjà utile.

Combinaisons possibles

Si l'on s'en tient à la combinatoire, le modèle paraît complexe. Il comporte en effet trois instances de décision possibles, trois degrés de connaissance de la catégorie d'appartenance, et trois possibilités pour prédire l'utilité, ce qui déterminerait 27 cas différents.

Simplifications

En fait, la majorité de ces cas peuvent être laissés de côté, ce qui simplifie grandement le problème.

Le présent texte, dans les pages précédentes, a déjà argumenté en faveur d'une seule échelle d'utilité combinée, et plusieurs des neuf cas restants peuvent encore être négligés.

Ainsi, le cas où l'estimation d'utilité ne serait possible que par catégorie, mais où l'on est sans information sur l'appartenance est naturellement un cas limite sans intérêt. On verra aussi que d'autres cas se recouvrent, au sens qu'ils appellent la même démarche de classement.

Méthodes de placement

1) Appartenance connue/ Utilité estimée individuellement hors catégories: Classement différentiel

C'est le cas d'information maximale. Il correspond au modèle classique de la psychométrie, où la population d'ensemble des candidats est considérée comme homogène et où la même équation de régression est appliquée à tous, pour estimer l'utilité attendue dans une profession. D'autres équations de régression sont naturellement utilisées pour chacune des autres orientations envisagées.

L'utilité maximale est atteinte pour l'ensemble des candidats à orienter si chaque personne est placée là où l'utilité est la plus grande *pour cette personne*. C'est ce qu'on appelle le classement différentiel.

Bien que cette stratégie paraisse relativement évidente, notons qu'elle ne correspond pas du tout aux méthodes de recrutement des écoles et des entreprises.

Ces dernières effectuent en effet un classement de leurs candidats, *en les comparant les uns aux autres* et en choisissant ceux qui ont le pronostic d'utilité le plus élevé.

L'école ou l'entreprise qui choisit en premier se réserve les candidats du niveau général le plus élevé. Les organisations qui sélectionnent ensuite n'ont plus le choix qu'entre les candidats qui restent, de niveau général nécessairement plus bas.

On peut appeler cette stratégie le placement par sélections successives. Elle se caractérise également par le fait que l'utilité de l'orientation n'est prise en compte que pour les candidats sélectionnés, et non pour les candidats rejetés.

Le classement différentiel conduit au contraire à considérer en même temps toutes les orientations possibles et à déterminer un placement optimal pour l'ensemble des candidats.

Aucune orientation n'est alors privilégiée par rapport aux autres (sauf si l'on introduit des coefficients d'importance). Des candidats de niveaux bas comme de niveau élevé se trouvent donc orientés dans chacune des voies envisagées, l'élément déterminant du choix étant que, pour le candidat considéré, la voie choisie est la meilleure.

La solution optimale est un peu plus difficile à déterminer si des quotas différents sont à satisfaire selon les orientations. Cette situation sera examinée plus bas.

Le principe général restera cependant valable, de comparer les points du profil pour chaque personne (d'où le qualificatif de "différenciel") et non de comparer les candidats entre eux.

2) Appartenance connue / Utilité estimée individuellement par catégorie: Classement modéré

L'utilité de placer chaque personne dans chaque orientation peut aussi être connue si la population est constituée de groupes hétérogènes, à condition que les appartenances soient connues et que des pronostics d'utilité soient calculables séparément pour chaque groupe.

On peut avoir par exemple une équation de régression pour les élèves réguliers et une autre pour redoublants, ceci pour chacune des orientations scolaires considérées. On ne tiendra compte, naturellement, que du pronostic approprié, correspondant au passé scolaire de chaque élève.

On appelle ce type de classement "modéré" parce que l'information sur l'appartenance sert de variable modératrice, conduisant à choisir la fonction prédictrice la plus appropriée.

Par exemple, comme il a été dit plus haut, les grandes universités américaines ont enregistré la carrière au sein de leurs Facultés des élèves venus des principales écoles secondaires du pays. Ceci leur permet d'apprécier plus exactement les chances de succès des nouveaux candidats, sur la base d'équations de régression appropriées à chaque école d'origine.

3) Appartenance connue / Utilité estimée globalement par catégorie: Classement par programmation linéaire

Une telle utilisation d'équations de régression pour estimer des utilités individuelles n'est réalisable, c'est certain, que dans de grandes organisations capables de mener de longs travaux de recherche. Dans beaucoup d'autres cas, il faut se contenter d'estimations globales de coûts, valables pour toute une catégorie de personnes. Ceci peut s'appliquer à la minimisation des coûts de formation, comme dans l'exemple (imaginaire) suivant.

Une entreprise a besoin de former 20 tourneurs et 25 fraiseurs professionnels. Elle veut les trouver au sein de son personnel, où sont disponibles 10 tourneurs spécialisés, 15

opérateurs de tours automatiques et 20 manoeuvres. Le temps de formation complémentaire nécessaire est donné dans le tableau 1, pour chaque catégorie.

La solution optimale se trouve facilement ici, en orientant vers le tournage (en fonction de la différence des temps de formation) les tourneurs spécialisés et 10 des 15 opérateurs. On arrive ainsi à un total de 3650 jours de formation, au lieu des 3850 qui auraient été nécessaires si l'on avait orienté les manoeuvres vers le tournage, ce qui était pourtant le plus facile pour eux.

De telles situations sont à traiter par des logiciels de programmation linéaire, dès que leur complexité est supérieure à celle de cet exemple.

Tableau 1.
Temps de formation complémentaire représentant les coûts
à minimiser, en respectant les contraintes marginales

Profession de départ	Profession d'arrivée		Nombres donnés
	Tourneur	Fraiseur	
Tourneur spécialisé	50 jours	80 jours	10 personnes
Opérateur de tour automat.	70 jours	90 jours	15 personnes
Manoeuvre	85 jours	100 jours	20 personnes
Nombres requis	20 personnes	25 personnes	45 personnes

4) Appartenance probable / Utilité estimée individuellement hors catégorie: Classement probabiliste, cas particulier du classement combiné

Dans bien des cas, l'appartenance à une catégorie ne peut être connue qu'en probabilité.

On peut imaginer que c'est le cas pour le Service chargé de répartir les étudiants en médecine entre diverses Facultés, de manière à utiliser au mieux les places de formation dans les hôpitaux.

Chaque étudiant a une probabilité différente d'abandonner ses études selon le lieu où il sera affecté. Ces probabilités d'appartenir à la catégorie "Abandon" peuvent être estimées en fonction de plusieurs paramètres, dont par exemple la distance de chaque hôpital par rapport au domicile primitif de l'étudiant.

Ce dernier a d'autre part une probabilité d'échec qui est, elle, fonction de ses capacités et des exigences de la Faculté à laquelle on envisage de l'affecter. (L'estimation des deux probabilités peut être obtenue à l'aide de la loi multinormale, si l'on connaît la position de l'étudiant dans l'espace des prédicteurs).

En admettant que ces deux probabilités soient indépendantes, parce que résultant de facteurs très différents, la "probabilité globale de départ" d'un étudiant dans une Faculté est le produit des deux probabilités précédentes. C'est la façon proposée par Tatsuoka (1956) pour opérationnaliser l'utilité attendue de chaque affectation pour chaque étudiant.

Elle suppose que les "départs" dûs à un abandon et ceux dûs à un échec ont la même gravité (utilité négative).

On cherchera naturellement à choisir (par programmation linéaire) les affectations des étudiants de façon à minimiser la somme des probabilités globales de départ. On peut appeler cette façon de faire le classement probabiliste.

L'élégance de la procédure précédente tient au fait que deux catégories seulement sont envisagées, les partants et les restants, de sorte que le calcul effectué sur les partants aboutit au même résultat que si l'on s'était intéressé aux restants. Il n'a donc pas à être répété pour l'estimation d'une utilité moyenne attendue de tous les cas possibles.

5) Appartenance probable / Utilité estimée individuellement par catégorie: Classement selon les risques connaissables: cas particulier du classement combiné

Imaginons qu'à un certain moment de la scolarité les élèves doivent s'orienter entre trois écoles: technique, commerciale et de culture générale. Des équations de régression, établies à l'intérieur de chaque école, permettent de pronostiquer le niveau de réussite en fin d'études, mais comme elles ont été établies sur des populations très différentes, on craint qu'elles ne s'appliquent pas telles quelles à la population générale.

On peut cependant parvenir à un pronostic raisonnable, à condition de procéder en deux étapes. Dans une première phase, on calcule, à partir des résultats d'un élève, quelle est sa probabilité de faire partie de chacun des trois groupes envisagés. Il suffit de connaître la distribution multinormale des membres de chaque groupe pour pouvoir calculer la densité de chaque groupe au point où se situe l'élève. On obtient les probabilités d'appartenance à posteriori en multipliant ces densités par l'importance numérique relative des trois groupes. On peut conclure cette étape en disant par exemple que sur 100 élèves ayant les mêmes caractéristiques que celui que l'on examine, 25 choisissent normalement l'école technique, 35 l'école commerciale et 45 l'école de culture générale.

Dans la seconde phase, on calcule les trois niveaux de réussite en fin d'études attendus d'un élève ayant ce profil de résultats, dans l'hypothèse où il ferait partie de chacun des trois groupes.

On peut alors raisonner de la façon suivante. Si l'on orientait les 100 élèves vers l'école technique, on connaît le résultat attendu pour 25 d'entre eux, mais on sait aussi que l'on en empêcherait 35 d'avoir le résultat prévu dans une école commerciale, et 45 de même dans une école de culture générale. On peut donc faire le total des résultats positifs et négatifs ainsi attendus, pour chacune des trois orientations, et classer les élèves de façon à maximiser la somme de ces valeurs. Il est facile de démontrer qu'on obtiendrait le même classement en calculant seulement pour chaque orientation le produit du résultat attendu par la probabilité a posteriori d'appartenir à cette orientation.

Cette façon de faire a l'avantage de ne faire appel qu'à des coûts connus, même si elle encourt certains risques en ne tenant pas compte de l'utilité possible de l'orientation d'un élève hors de sa catégorie d'appartenance.

6) Appartenance probable / Utilité estimée globalement par catégorie: Classement par fonction discriminante pondérée

Dans les cas où un pronostic individuel d'utilité n'est pas calculable, il est pourtant souvent possible d'estimer approximativement le coût des erreurs de classement et d'en tirer un indice pouvant guider l'affectation. On doit pour cela partir d'une matrice donnant le coût à attendre si l'on affecte à la catégorie "colonne" quelqu'un qui appartiendrait effectivement à la catégorie "ligne". Un exemple en est donné au tableau 2.

Tableau 2.
Exemple imaginaire de coûts d'erreurs de classement

Appartenance réelle	Décision d'affectation		
	Tourneur	Ajusteur	Serrurier
Tourneur	0	10	5
Ajusteur	2	0	3
Serrurier	8	5	0

Dans un contexte industriel, on peut considérer que les demandes de déplacement, les insuccès et les départs révèlent un placement défectueux. On peut alors chiffrer approximativement le coût d'une nouvelle affectation. Par exemple, il semble coûteux d'après le tableau 2 d'affecter par erreur comme ajusteur quelqu'un qui était en fait tourneur.

Si l'on connaît la distribution multinormale de chaque groupe, on peut calculer quelle est la densité de chacun au point représentatif du sujet à orienter. Ces probabilités, une fois multipliées par la probabilité a priori de chaque groupe et divisées par la densité totale au point considéré, donnent les probabilités a posteriori que la personne en question provienne de chacun des groupes.

Le coût attendu des erreurs de placement comme tourneur sera alors 2 fois la probabilité a posteriori que l'intéressé soit ajusteur, plus 8 fois la probabilité qu'il soit serrurier. Le coût des erreurs de placement comme ajusteur sera 10 fois la probabilité comme tourneur plus 5 fois la probabilité comme serrurier, et similairement pour la dernière colonne de coûts. Si les quotas sont libres, on peut proposer l'orientation pour laquelle le coût attendu est le moins élevé.

7) Appartenance inconnue / Utilité estimée globalement par catégorie: Classement minimax

Si l'on énumère tous les cas possibles, le 7ème devrait correspondre à une appartenance inconnue, mais à une possibilité d'estimer individuellement l'utilité attendue de chaque orientation. Ce cas se ramène au premier, celui auquel s'applique le classement différentiel, qui ne prend pas en compte les probabilités d'appartenance.

La 8ème combinaison est impossible. On ne peut pas estimer individuellement une utilité dépendant de la catégorie lorsque la catégorie ne peut pas être connue. Ce cas n'est donc pas à considérer non plus.

La 9ème combinaison ressemble assez à la 6ème. Elle suppose qu'une matrice de coûts est à disposition pour pondérer chaque type d'erreur de classement, comme dans le tableau 2. La difficulté est cependant que, cette fois, aucune appartenance n'est plus probable qu'une autre et que le choix d'orientation paraît alors purement arbitraire.

On pourrait envisager cependant d'appliquer une stratégie minimax, qui représente dans la théorie des jeux une solution pessimiste, mais prudente. Lorsqu'on doit choisir entre divers coups, face à un joueur adverse qui est prêt à exploiter au maximum la situation contre nous, la stratégie minimax consiste à choisir l'option la moins mauvaise, celle dont l'adversaire pourra tirer le plus faible bénéfice. Toute autre option risquerait de lui offrir l'occasion de gains plus élevés encore. Par exemple, avec la matrice de pertes du tableau 2, l'affectation recommandée serait la dernière colonne, parce que c'est pour elle que le coût le plus grand qu'on puisse craindre est minimum, soit 5, (alors qu'il peut être de 8 et de 10 pour les autres options).

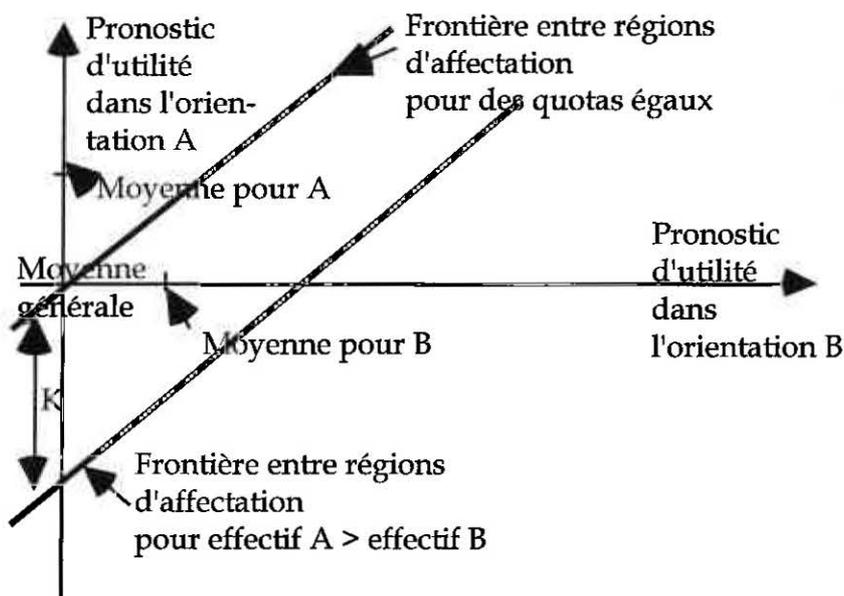
Dans une situation d'orientation, cette stratégie signifierait que tout ce qui peut arriver de mauvais arrivera, et que l'on se prépare seulement à éviter le pire. C'est là une attitude pessimiste, acceptable sans doute dans certaines situations, pour une politique de prévention des accidents par exemple, mais qu'on ne peut guère recommander comme stratégie d'orientation.

Importance des quotas

C'est pour ne pas alourdir la présentation des différents types de classement que la différence n'a pas été davantage soulignée ci-dessus, entre situations à quotas libres et à quotas imposés. Elle est pourtant fondamentale, car les difficultés n'apparaissent en fait qu'avec des quotas imposés. Sinon, il suffit de laisser chaque candidat s'orienter vers la voie d'utilité maximale pour lui.

On peut représenter par un graphique la solution optimale dans l'hypothèse de quotas libres. Elle apparaît à la Figure 1, pour le cas simple où seulement deux orientations sont considérées. Les échelles d'utilité des orientations A et B sont les axes par rapport auxquels chaque candidat est situé. L'ensemble des candidats forme alors un nuage de points, en forme d'ellipse centrée sur l'origine, du fait de la corrélation existant habituellement entre les pronostics d'utilité.

Figure 1. Représentation du classement différentiel en 2 dimensions



Pour des quotas libres, chaque personne est affectée du côté où son utilité est la plus grande. Géométriquement, c'est la bissectrice de l'angle formé par les deux axes qui détermine la frontière, puisque c'est le lieu des points où la projection sur l'axe A est égale à la projection sur l'axe B.

Le classement optimum dans le cas d'effectifs inégaux est déterminé par la position de la seconde ligne oblique. Cette frontière entre les deux régions (d'attribution à l'orientation A ou B) est le lieu des points comportant une différence égale à K entre les deux pronostics. C'est en faisant varier K que l'on admet plus de candidats dans une région ou dans l'autre. Lorsque $K = 0$, une proportion égale de candidats est admise dans les deux orientations.

Le gain d'utilité dû à cette orientation optimale apparaît dans le fait que la moyenne des pronostics d'utilité, aussi bien pour A que pour B, est supérieure à la moyenne générale.

Avec trois orientations ou plus, ce sont encore les différences de pronostics d'utilité qui détermineront le choix de l'orientation. On fera varier les quotas admis dans chaque voie en modifiant les valeurs des différents K_{ij} pour chaque paire (i,j) d'orientations.

La même technique de répartition pourrait être appliquée à toutes les valeurs d'utilité dont il a été question plus haut, lorsque les quotas sont différents. Naturellement, avec d'avantages d'orientations, le calcul des différents K_{ij} et la répartition se ferait par ordinateur.

Applications possibles

Possibilités du modèle

Il est évident que ce modèle n'est pas "applicable" tel quel, mais il n'est pas nécessaire de le concrétiser pour en tirer néanmoins des indications utiles.

Principe de décision

A la base du modèle, on trouve en effet un principe qu'il est important d'explicitier. On raisonne le choix en prenant l'attitude non pas du joueur attiré par l'importance du "gros lot", ni du pessimiste qui n'envisage que le pire, mais de l'actuaire qui examine la valeur attendue en probabilité de chacune des options possibles.

Utilités à la fois individuelles et sociales

Le conseil d'orientation ne peut se contenter des seules valeurs individuelles ou institutionnelles, mais doit chercher à intégrer dans la réflexion des considérations sociales, tout en donnant la priorité à la personne.

Orientation plutôt que sélection

On a vu qu'une répartition par sélections successives devait différencier le plus possible les candidats les uns par rapport aux autres, et par conséquent s'appuyer sur la dimension du facteur général pour les classer sans ambiguïté. La hiérarchisation des orientations qui s'ensuivait avait de graves conséquences psychologiques et sociales.

Au contraire une véritable orientation, par opposition à ces sélections successives, ne privilégie pas certaines voies par rapport à d'autres, assure des distributions de compétences plus égales entre professions, et s'appuie sur les différences de pronostic à l'intérieur d'une même personne et non sur les différences entre personnes. Elle s'intéresse pour cela moins au facteur général qu'aux facteurs de compétence spécifiques qui peuvent justifier le choix d'une orientation plutôt que d'une autre. Elle obtient ainsi une meilleure utilisation des compétences de l'ensemble de la population.

Possibilité de tenir compte à la fois des intérêts et des aptitudes

On a vu que certains types de classement combinaient probabilités d'appartenance et pronostics d'utilité. Or les appartenances sont surtout fonction des intérêts et des motivations, tandis que la réussite dépend surtout des aptitudes.

Ceci suggère la possibilité d'une combinaison, au moins intuitive, des deux types d'information.

Nécessité de tenir compte à la fois des traits *typiques* et *prédicteurs*

Il ne faudrait pas s'en tenir à une approche unique dans la logique du conseil d'orientation: (à quel type professionnel la personne ressemble-t-elle? ou au contraire, dans quelle profession réussira-t-elle?). Il est possible de combiner ces deux logiques dans le cadre de la théorie de la décision.

Nécessité d'un cadre de référence commun

En rapport avec cette volonté de faire la synthèse des informations disponibles, on peut souligner la nécessité de tenir compte à la fois des différences intergroupes et intragroupes. Cela suppose l'application d'une même batterie d'instruments à tous les candidats, quel que soit leur niveau, ou du moins cela implique leur positionnement sur des échelles de performance communes, en utilisant la théorie de réponse aux items pour effectuer un examen sur mesure.

Développements possibles

Une revue de la littérature et une analyse plus détaillée des différents types de classement a été rédigée. Elle est disponible sous forme miméographiée. (*La répartition du personnel* - S'adresser à l'auteur). Un abrégé en a été publié dans le *Traité de Psychologie Appliquée* de Reuchlin, volume IV.

Limites du modèle

Le modèle ne prétend pas dicter la démarche du conseiller d'orientation. Les problèmes de l'orientation se situent dans un monde beaucoup plus complexe que celui des abstractions prises en compte dans les pages précédentes et il ne faudrait pas limiter son attention aux trois facettes envisagées... mais cela va sans dire et le lecteur l'aura bien compris!

RÉFÉRENCES

- Cardinet, J. (1969). *La répartition du personnel*. Neuchâtel : Université, Institut de psychologie (miméogr.), 107 p.
- Cardinet, J. (1971). La mise en place des hommes. In M. Reuchlin (Ed.), *Traité de psychologie appliquée. Vol. IV, ch. 3* (pp. 111-167). Paris : PUF.
- Cronbach, L.J. & Gleser, G.C. (1965). *Psychological tests and personnel decisions*. Urbana : University of Illinois Press.
- Ghiselli, E.E. (1960). The prediction of predictability. *Educ. psychol. Measmt.*, 20, 3-8.
- Tatsuoka, M.M. (1956). *Joint probability of membership in a group and success therein : an index which combines the information from discriminant and regression analysis*. Cambridge, Mass. : Harvard Graduate School of Education (Unpublished doctoral dissertation).