

REVUE DE STATISTIQUE APPLIQUÉE

B. IVANOVIC

Comment établir une liste des indicateurs de développement

Revue de statistique appliquée, tome 22, n° 2 (1974), p. 37-50

http://www.numdam.org/item?id=RSA_1974__22_2_37_0

© Société française de statistique, 1974, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

COMMENT ÉTABLIR UNE LISTE DES INDICATEURS DE DÉVELOPPEMENT

B. IVANOVIC

1 – INTRODUCTION

Le choix des indicateurs de développement socio-économique est certainement un des plus importants problèmes dans l'exécution du projet d'évaluation et de prévision du développement durant la Deuxième Décennie des Nations Unies.

La qualité des résultats sera d'autant meilleure que l'information sur le niveau du développement des pays, contenue dans les indicateurs choisis, est plus complète.

Il existe aujourd'hui un certain nombre d'indicateurs qui sont bien enracinés dans notre esprit comme étant ceux qui renseignent le plus sûrement sur le développement socio-économique d'un pays et qui sont utilisés le plus souvent dans l'évaluation de son niveau de développement. Par exemple, des indicateurs comme le P.I.B. par habitant, le taux d'analphabètes, le pourcentage de la population non agricole ou la part des produits industriels dans le P.I.B. représentent quelques-uns de ces indicateurs classiques.

Cependant, les listes d'indicateurs de développement utilisées dans les différentes institutions de recherche, nationales ou internationales, ne sont pas toujours identiques et les controverses ne manquent pas dans la discussion sur l'utilité et l'importance de l'un ou de l'autre de ces indicateurs classiques.

En réalité, on peut se demander d'une part s'il existe des raisons objectives pour permettre de considérer tous ces indices comme des indicateurs de développement socio-économique et, d'autre part, si d'autres indices, ignorés ou négligés jusqu'ici, ne seraient pas plus appropriés que ne le sont certains de ceux que l'on envisage traditionnellement.

En recherchant ainsi de nouveaux indicateurs de développement, nous avons défini et étudié l'indicateur de la "Concentration des naissances d'après

l'âge de la mère”(1). C'est un indicateur qui donne l'information sur la qualité familiale d'un pays. Mais il n'est pas difficile de constater qu'il s'agit aussi d'un important indicateur de développement. Remarquons, tout d'abord, que la haute concentration des naissances dans un pays, calculée d'après l'âge de la mère, signifie que par famille le nombre des enfants est limité, ce qui assure un niveau de vie plus élevé dans ce pays. Ensuite, que le domaine de concentration se trouve entre 20 et 30 ans d'âge de la mère, ce qui veut dire que l'âge des parents permet aux enfants d'avoir une bonne éducation familiale, une protection économique prolongée, un développement physique favorable et une éducation scolaire plus complète. Enfin, si la concentration est grande, on rencontre rarement des cas défavorables comme ceux, par exemple, de mères trop jeunes ou trop âgées, d'enfants nés d'un second mariage, etc. En bref, la vie d'une famille est d'autant plus ordonnée et son activité sur le plan socio-économique est d'autant plus efficace que la concentration des naissances, d'après l'âge de la mère, est plus forte.

Les analyses quantitatives montrent que cet indicateur donne plus d'information sur le niveau de développement d'un pays que certains indicateurs classiques comme, par exemple, les exportations d'articles manufacturés par habitant, la durée de vie moyenne, le taux d'analphabètes la part des industries manufacturières dans le P.I.B., le taux d'inscriptions scolaires, etc.

On pourrait conclure que, pour arriver à une liste exhaustive des indicateurs de développement, il faudrait ajouter aux indicateurs de toutes les listes utilisées jusqu'à présent ceux qui ont été négligés ou ignorés dans le passé.

On formerait ainsi une très longue liste qu'on appellerait la liste maximale des indicateurs de développement socio-économique. Les statistiques de ces indicateurs donneraient des informations très complètes sur tous les aspects du développement des pays observés.

Cependant, en examinant toutes les publications statistiques du Bureau de statistique des Nations Unies, des commissions économiques régionales, des diverses agences spécialisées des Nations Unies ainsi que les publications nationales et d'autres documents publiés, soit officiellement soit officieusement, nous pouvons constater que le choix des indicateurs de développement est très limité du fait de l'absence de données statistiques relatives à de nombreux pays en voie de développement.

(1) Pour le pays j, la mesure de cet indicateur est donnée par

$$C_j = 1\,000 \sqrt{\frac{1}{n} \sum_{i=j}^n \left(\frac{x_{ij}}{x_j} - \frac{1}{n} \right)^2},$$

où n représente le nombre de groupes d'âge, x_{ij} le nombre de naissances où l'âge de la mère appartient au i-ème groupe et X_j le nombre total de naissances dans le pays j au cours de l'année observée.

L'absence ou la mauvaise qualité des données statistiques empêche l'introduction d'un grand nombre d'indicateurs de développement comme, par exemple, ceux qui renseignent sur la productivité du travail dans l'industrie ou dans l'agriculture, sur le niveau de vie de la population, l'infrastructure, les cadres dans l'industrie et l'enseignement, etc.

L'introduction d'un seul indicateur dans le système statistique d'un pays exige souvent l'organisation d'un nouveau service doté de cadres compétents et d'un équipement adéquat qui serait apte à compiler, traiter et publier les données. Ces données devraient être d'une bonne qualité et comparables à celles des autres pays.

Par conséquent, l'introduction de nouveaux indicateurs dans notre liste d'indicateurs de développement exige du temps, de nouvelles institutions et de grands sacrifices financiers. Il va sans dire que ce problème est encore plus grave dans les pays en voie de développement.

Pour suivre le développement des pays au cours de la Deuxième Décennie, avec une précision qui serait suffisante pour atteindre nos objectifs, quelques nouveaux indicateurs devraient incontestablement être introduits dans la liste d'indicateurs de développement. Mais lesquels faut-il choisir parmi tant d'indicateurs déjà proposés ?

Le premier problème à résoudre dans l'exécution du projet d'évaluation et de prévision du développement au cours de la Deuxième Décennie sera donc l'établissement scientifique d'une liste optimale d'un nombre réduit d'indicateurs de développement.

2 – CRITERE DU CHOIX DES INDICATEURS DE DEVELOPPEMENT SOCIO-ECONOMIQUE

Divers indicateurs statistiques peuvent être utilisés pour estimer et suivre le développement socio-économique d'un pays. Chacun de ces indicateurs apporte une information partielle sur le niveau du développement et, en général, ces informations partielles ne sont pas identiques. Ainsi, par exemple, l'indicateur "Pourcentage de la population non agricole" nous apporte une information sur le niveau du développement d'un pays qui n'est pas identique à celle apportée par l'indicateur "Consommation d'énergie par habitant". Il faut donc choisir et former tout un ensemble d'indicateurs pour avoir une idée plus complète sur le niveau du développement des pays observés.

Première conclusion : *En augmentant le nombre d'indicateurs, on augmente l'information globale sur le niveau de développement des pays observés.*

Les différents indicateurs de développement ne contiennent pas le même volume d'information sur le niveau de développement d'un pays. Cela veut dire qu'ils n'ont pas la même importance par rapport au développement. Ainsi,

“P.I.B. par habitant” apporte plus d’information sur le niveau du développement socio-économique d’un pays que, par exemple, l’indicateur “Durée de vie moyenne des habitants”. Si la quantité d’information contenue dans un indicateur de développement est mesurable, on pourrait établir un classement de ces indicateurs d’après leur importance.

Deuxième conclusion : *Des ensembles d’indicateurs de même taille (même nombre d’indicateurs) ne contiennent pas, en général, la même quantité d’information sur le niveau de développement d’un pays.*

Les indicateurs de développement ne sont, en général, pas indépendants les uns des autres de sorte que l’information fournie par un indicateur est contenue, en partie, dans l’information qui a déjà été fournie par les autres indicateurs. En prenant donc en considération tout l’ensemble des indicateurs choisis, il se pourrait que certaines quantités d’information se répètent deux ou plusieurs fois. Il va sans dire qu’il faudrait soustraire ces redoublements dans l’information globale et ne retenir que l’information propre contenue dans cet ensemble d’indicateurs.

Troisième conclusion : *L’information globale (sans redoublements) d’un ensemble d’indicateurs est, en général, inférieure à la somme des quantités d’information contenues individuellement dans chaque indicateur de cet ensemble.*

Comme exemple, prenons le cas de deux indicateurs dont la dépendance mutuelle est totale. Si nous rangeons les pays observés par rapport aux valeurs correspondantes de l’un ou de l’autre indicateur, nous obtiendrons deux classements identiques. Cela signifie que les deux indicateurs contiennent la même information sur le niveau de développement des pays observés et que, par conséquent, il suffit d’en prendre un seul, l’autre étant superflu.

Quatrième conclusion : *En dépit du fait que deux indicateurs observés peuvent être très importants au point de vue de l’information qu’ils donnent, séparément, sur le niveau de développement d’un pays, la contribution de l’un d’eux est insignifiante s’il existe une forte corrélation entre les deux.*

Cinquième conclusion : *Il faut supprimer les indicateurs dont la quantité d’information est entièrement contenue dans l’information globale des indicateurs déjà pris en considération.*

Il s’ensuit qu’un petit nombre relatif d’indicateurs, judicieusement choisis, pourrait contenir une plus grande information globale qu’un grand nombre d’indicateurs mal choisis.

Par conséquent, pour avoir une idée aussi complète que possible sur le niveau du développement des pays observés, il ne suffit pas d’augmenter le nombre des indicateurs. Pour augmenter l’information globale, il faut tenir compte de l’importance de chaque indicateur, exprimée sous la forme d’une quantité d’information, ainsi que des redoublements dans l’information, exprimés sous la forme des dépendances mutuelles entre les indicateurs.

Pour faire un choix optimal d'un nombre limité d'indicateurs de développement, la quantité d'information globale (sans redoublements) devrait être maximum tandis que la somme des redoublements devrait être minimum.

La contribution d'un indicateur dans l'évaluation du niveau de développement d'un pays est d'autant plus élevée que la quantité d'information, non contenue dans l'information globale des indicateurs plus importants que lui, est grande.

Cependant, même s'il s'agit d'un indicateur dont la contribution propre est non négligeable, il ne nous sera guère utile si sa valeur ne varie presque pas dans l'ensemble des pays observés. Ainsi, par exemple, l'indicateur "Pourcentage d'analphabètes dans la population âgée de 10 ans ou plus" n'a aucune signification dans l'ensemble des pays riches où sa valeur est stabilisée entre 1 % et 3 %. Par contre, dans l'ensemble des pays pauvres, cette valeur varie entre 5 % et 99 %, de sorte que ce même indicateur est très discriminatif lorsqu'il s'agit d'évaluer le niveau de développement des pays en voie de développement. La situation serait inverse si nous prenions la "Part de la production industrielle dans le P.I.B." comme indicateur de développement.

Sixième conclusion : *Un bon indicateur de développement devrait être discriminatif dans l'ensemble des pays observés.*

Ainsi, nous pouvons compléter le critère de choix d'un indicateur de développement pour un ensemble de pays observés, de la façon suivante :

La contribution d'un indicateur dans l'évaluation du niveau de développement d'un pays est d'autant plus élevée que sa puissance de discrimination est forte dans l'ensemble des pays observés et que sa quantité d'information, non contenue dans l'information globale des indicateurs qui le précèdent et qui sont déjà pris en considération, est grande.

Remarquons que chacune de ces deux conditions n'est que nécessaire et nous acceptons leur conjonction comme suffisante.

3 – EVALUATION OBJECTIVE DU DEGRE D'IMPORTANCE D'UN INDICATEUR DE DEVELOPPEMENT

Le degré d'importance ou de signification d'un indicateur de développement pourrait être conçu comme la quantité d'information sur le niveau de développement d'un pays, contenue dans cet indicateur. Il va sans dire qu'il n'est pas question d'exprimer d'une manière absolue cette quantité d'information. Ce que nous pouvons déterminer, c'est plutôt une quantité relative qui dépendrait de l'ensemble de tous les pays observés et de l'ensemble des autres indicateurs choisis.

S'il existe une dépendance totale entre l'indicateur observé et le niveau de développement d'un pays, nous aurons le même classement des pays en

les rangeant soit par rapport à leur valeur de cet indicateur soit par rapport à leur niveau de développement. Nous pouvons dire, dans ce cas, que toute l'information sur le niveau de développement des pays est contenue dans ce seul indicateur, car il suffit d'établir une liste de classification des pays observés par rapport aux valeurs correspondantes de cet indicateur pour obtenir, en même temps, la liste de classification de ces pays par rapport à leur niveau de développement.

Si la dépendance est partielle, l'information contenue dans l'indicateur observé n'est qu'une partie de l'information totale sur le niveau de développement des pays.

Il s'ensuit que *nous pouvons admettre que le degré de l'importance d'un indicateur de développement correspond au degré de dépendance entre cet indicateur et le niveau de développement.*

Cependant, la définition de la notion de développement socio-économique d'un pays n'est pas encore précisée d'une façon satisfaisante et il n'est pas possible de construire un indice global qui pourrait exprimer, d'une manière absolue, le niveau de ce développement. Nous ne pouvons que nous contenter de déterminer, à l'aide des données d'un ensemble d'indicateurs choisis, la position relative d'un pays par rapport aux autres dans l'ensemble des pays observés. C'est ainsi que nous arrivons à la notion de "distance" entre deux pays par rapport à leur développement.

Cette distance devrait satisfaire les conditions que les propriétés de la distance d'un espace métrique imposent, celles qui assurent l'unicité de solution et enfin celles qui l'adaptent à la nature du problème, c'est-à-dire celles qui imposent la soustraction des redoublements dans l'information et qui diminuent les effets de corrélations apparentes.

En se basant sur un groupe d'indicateurs $X = \{X_1, \dots, X_n\}$, le niveau relatif du développement d'un pays sera alors défini comme distance dans le développement entre ce pays et un pays fictif auquel sont attribuées, pour chaque indicateur, des valeurs minima dans l'ensemble des pays observés $P = \{P_1, \dots, P_m\}$.

Ainsi, en utilisant la I-distance⁽¹⁾, la mesure relative du niveau de développement du pays P_r serait définie par l'expression suivante :

$$D(r) = \sum_{i=1}^n \frac{x_{ir} - x_i}{\sigma_i} \prod_{j=1}^{i-1} (1 - r_{ji.1 \dots j-1}), \quad (3.1)$$

où x_{ir} désigne la valeur de l'indicateur X_i dans le pays P_r , $x_i = \text{Min}\{x_{i1}, \dots, x_{im}\}$, σ_i l'écart-type de X_i et $r_{ji.1 \dots j-1}$ le coefficient de corrélation partielle entre X_i et X_j .

 (1) Voir le Research Memorandum N° 41 du 5 novembre 1970 (CNUCED).

Si tous les indicateurs choisis sont mutuellement indépendants, la distance D se réduit à la distance F de Fréchet, c'est-à-dire

$$F(r) = \sum_{i=1}^n \frac{x_{ir} - \bar{x}_i}{\sigma_i} \quad (3.2)$$

Parmi tous les indicateurs du développement, le P.I.B. par habitant contient indiscutablement la plus grande quantité d'information sur le développement d'un pays. Même les économistes considèrent, en général, cette quantité d'information comme suffisante pour avoir la première approximation du niveau de développement d'un pays.

D'autre part, l'ensemble des indicateurs de développement est l'union de deux grands groupes d'indicateurs : le premier englobant des indicateurs économiques et le deuxième des indicateurs sociaux. Tous les indicateurs dépendent du P.I.B. par habitant, mais la dépendance entre les indicateurs de différents groupes pourrait être seulement apparente. C'est pour cette raison qu'il conviendrait mieux de prendre en considération séparément le P.I.B. par habitant, le groupe des indicateurs économiques $\{X = X_1, \dots, X_n\}$ et le groupe des indicateurs sociaux $\{Y = Y_1, \dots, Y_n\}$. Dans ce cas, la mesure relative du niveau de développement du pays P_r prend la forme suivante

$$D(r) = \frac{d_0(r)}{\sigma_0} + \sum_{i=1}^{n'} \frac{d_i(r)}{\sigma_i} (1 - r_{oi}) \prod_{j=1}^{i-1} (1 - r_{ji \dots j-1}) \\ + (1 - R_{xy}) \cdot \sum_{i=1}^{n'} \frac{\Delta_i(r)}{s_i} (1 - \rho_{oi}) \prod_{j=1}^{i-1} (1 - \rho_{ji \dots j-1}), \quad (3.3)$$

où nous avons désigné par

- $d_i(r)$ — la différence entre les valeurs X_i des pays P_r et P_-
- $\Delta_i(r)$ — la différence entre les valeurs Y_i des pays P_r et P_-
- σ_i — l'écart-type de X_i
- s_i — l'écart-type de Y_i
- r_{oi} — le coefficient de corrélation simple entre X_i et le P.I.B. par habitant
- ρ_{oi} — le coefficient de corrélation simple entre Y_i et le P.I.B. par habitant
- $r_{ji \dots j-1}$ — le coefficient de corrélation partielle entre X_i et X_j
- $\rho_{ij \dots j-1}$ — le coefficient de corrélation partielle entre Y_i et Y_j
- R_{xy} — le coefficient de corrélation collective entre les groupes X et Y
— (ce coefficient exige le même nombre d'indicateurs de chaque groupe)

En prenant donc la I-distance D comme mesure relative du niveau de développement d'un pays et en exprimant le degré de dépendance entre l'indi-

icateur X_i et le niveau de développement par le coefficient de corrélation simple r_i , cette valeur r_i sera considérée comme degré de l'importance de X_i .

En rangeant les indicateurs observés d'après l'ordre de grandeur de leurs coefficients de corrélation simple avec la I-distance, nous obtiendrons le classement des indicateurs de développement d'après leur importance.

On pourrait ainsi classer objectivement les indicateurs de développement, d'après la corrélation entre les différents indicateurs effectivement utilisés et l'indice global contenant le maximum d'information. Cependant, le calcul des coefficients de corrélation ne peut pas être effectué directement puisqu'au moment où nous devrions calculer la valeur de ces coefficients, nous ne disposons pas encore de la série des I-distances dont le calcul nécessite la connaissance du classement des indicateurs que nous ne sommes qu'en train de rechercher.

En se trouvant ainsi dans un cercle fermé, nous nous demanderons s'il est au moins possible de résoudre le problème du classement des indicateurs de développement en appliquant la méthode d'itération approximative.

Remarquons, tout d'abord, que dans notre problème peuvent apparaître deux cas extrêmes. Dans le premier cas, tous les indicateurs sont indépendants et le problème de la multiplication des mêmes quantités d'information ne se pose plus. Dans l'autre la dépendance de tous les indicateurs choisis par rapport à l'un d'eux est totale. Les quantités d'information de tous les indicateurs sont alors contenues dans celui de l'indicateur dominant. Dans les deux cas, on peut sans aucune difficulté établir la série des I-distances et déterminer ensuite l'ordre des indicateurs par rapport à leur degré d'importance.

Cependant, dans la pratique, on se trouve généralement entre ces deux extrémités. Mais on peut dire qu'on se trouve toujours soit dans la situation où un indicateur contient la plus grande partie de l'information globale sur le niveau de développement des pays, soit dans la situation où l'on ne peut pas affirmer qu'un indicateur est, d'une façon évidente, dominant par rapport aux autres.

Le premier cas est plus proche du cas extrême de la dépendance totale de tous les indicateurs par rapport à l'indicateur dominant que du cas extrême de l'indépendance totale de tous les indicateurs. C'est pourquoi, en première approximation, nous supposons que l'indicateur dominant contient la totalité de l'information globale existant dans l'ensemble de tous les indicateurs observés. En d'autres termes, nous supposons provisoirement que le niveau de développement d'un pays correspond à la valeur de l'indicateur dominant pour ce pays.

En calculant la matrice de corrélation et en rangeant les indicateurs observés d'après l'ordre de grandeur de leurs coefficients de corrélation avec l'indicateur dominant, nous obtiendrons un classement provisoire des indicateurs de développement d'après leur degré d'importance. On dira qu'il s'agit du classement de la première étape.

En utilisant donc, provisoirement, cet ordre d'indicateurs, nous pouvons calculer les valeurs de la I-distance pour tous les pays considérés. Etant donné que cette I-distance, basée sur tous les indicateurs observés, contient plus d'information sur le niveau de développement des pays que l'indicateur dominant, nous déterminerons, dans la deuxième étape, l'ordre des indicateurs d'après la grandeur de leurs coefficients de corrélation avec la I-distance de la première étape. Ce serait le classement de la deuxième étape.

En calculant la série des I-distances d'après le classement de la deuxième étape et en calculant les coefficients de corrélation entre les indicateurs observés et la I-distance ainsi obtenue, nous obtiendrons le classement des indicateurs de la troisième étape.

On continue ce procédé itératif jusqu'au moment où le classement des indicateurs est identique avec celui de l'étape précédente. Ce classement d'indicateurs est définitif car il reste invariable dans toutes les étapes postérieures.

Dans le deuxième cas (sans indicateur dominant), on est plus proche du cas extrême de l'indépendance totale de tous les indicateurs que du cas extrême de la dépendance totale de tous les indicateurs par rapport à l'indicateur dominant. Alors nous supposons, dans la première approximation, que tous les indicateurs sont indépendants et que, par conséquent, il n'y a pas de redoublements dans l'information.

La I-distance se réduit dans ce cas à la distance F de Fréchet dont le calcul n'exige aucun ordre des indicateurs puisqu'ils entrent avec le même droit dans la formule de F.

En calculant les valeurs de la distance F pour tous les pays considérés et en déterminant ensuite les coefficients de corrélation entre les indicateurs observés et la distance F, nous obtiendrons le classement de la première étape des indicateurs de développement d'après leur degré d'importance.

Par la suite, on utilise la même méthode itérative, comme dans le premier cas, pour arriver au classement définitif des indicateurs.

Remarquons que dans le problème du choix des indicateurs du développement socio-économique on se trouve dans le premier cas, l'indicateur "P.I.B. par habitant" étant, sans aucun doute, dominant.

Si les indicateurs de développement observés $\{X_1, X_2, \dots, X_n\}$ sont rangés d'après le classement définitif et si la I-distance D pour chaque pays est calculée d'après ce classement, la mesure quantitative de l'importance de l'indicateur X_i sera donnée par la formule

$$I_i = r(X_i ; D).$$

Leur ordre de grandeur correspond au classement définitif, c'est-à-dire :

$$r(X_1 ; D) \geq r(X_2 ; D) \geq \dots \geq r(X_n ; D).$$

En prenant le groupe des pays en voie de développement et en se limitant à 12 indicateurs de développement, en 1968, l'ordre d'importance de ces indicateurs était le suivant :

Ordre	Indicateur	Degré d'importance
1	Le P.I.B. par habitant	0,97
2	La consommation d'énergie par habitant	0,95
3	Le P.I.B. de la population active dans l'agriculture	0,81
4	Le nombre de médecins par 100 000 habitants	0,81
5	Le tirage de journaux par 1 000 habitants	0,80
6	Le pourcentage de la population non agricole	0,79
7	La concentration des naissances d'après l'âge de la mère	0,77
8	Les exportations d'articles manufacturés par habitant	0,76
9	La durée de vie moyenne	0,71
10	Le taux d'alphabètes	0,69
11	La part des industries manufacturières dans le P.I.B.	0,66
12	Le taux d'inscriptions scolaires	0,64

D'après ces résultats, on constate que les indicateurs, tels que le taux d'alphabètes et la part des industries manufacturières dans le P.I.B., sont parmi les plus faibles informateurs sur le niveau de développement des pays dans l'ensemble des indicateurs observés. Cependant, mis à part le P.I.B. par habitant, c'est précisément ces deux indicateurs que le Comité de la planification du développement des Nations Unies a choisis pour identifier le groupe des pays les moins avancés.

4. PUISSANCE DISCRIMINATIVE D'UN INDICATEUR

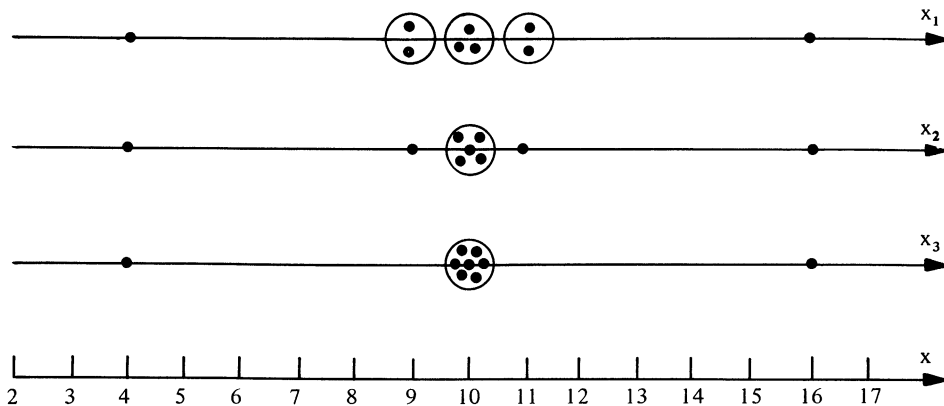
Comme mesure de la puissance discriminative de l'indicateur X_i dans l'ensemble des pays P , nous proposons le coefficient de discrimination dont l'expression algébrique aurait la forme suivante

$$CD_i(P) = \frac{2}{m(m-1)} \sum_{j, i > j}^{k_i} m_{ij} m_{i1} \left| \frac{x_{ij} - x_{i1}}{x_i} \right| \quad (4.1)$$

où m représente le nombre de pays dans l'ensemble P , k_i le nombre de différentes valeurs, que X_i prend dans l'ensemble P et m_{ij} le nombre de pays dont la valeur de X_i est x_{ij} .

Remarquons que ce coefficient n'a pas la même signification que le coefficient de variation, qui mesure la dispersion moyenne relative de X_i dans P .

Ainsi, dans notre exemple illustratif de trois répartitions où la diversité de la puissance de discrimination est évidente (avec les coefficients de discrimination respectifs 0.322, 0.300 et 0.267), le coefficient de variation est presque constant (0.288, 0.286 et 0.283).



5. – INFORMATION GLOBALE D'UN GROUPE D'INDICATEURS DE DEVELOPPEMENT

Afin de déterminer la quantité de l'information globale, contenue dans l'ensemble de k indicateurs choisis X_1, X_2, \dots, X_k , il faut s'en tenir au critère relatif au choix des indicateurs de développement qui a été formulé dans le paragraphe 2 et qui nous impose de tenir compte du degré d'importance de chaque indicateur, des redoublements dans l'information contenue dans les différents indicateurs et de la puissance discriminative de chaque indicateur.

Supposons que les indicateurs $\{X_1, X_2, \dots, X_k\}$ soient déjà ordonnés d'après leur degré d'importance.

La contribution globale de cet ensemble d'indicateurs dans l'évaluation du niveau de développement des pays observés est donnée par l'expression

$$C(1, 2, \dots, k) = \sum_{i=1}^k CD_i(P) \prod_{j=1}^{i-1} (1 - r_{ji.12\dots j-1}), \quad (5.1)$$

où $CD_i(P)$ représente le coefficient de discrimination de l'indicateur X_i dans l'ensemble des pays P et $r_{ji.12\dots j-1}$ le coefficient de corrélation partielle entre les indicateurs X_i et X_j .

Cette expression est croissante et homogène par rapport aux puissances discriminatives des indicateurs choisis. Chaque puissance est pondérée de telle façon que la contribution de l'indicateur correspondant dans l'information globale est d'autant plus importante que le rang de l'indicateur est plus élevé et les corrélations partielles, entre cet indicateur et ceux qui lui précèdent, plus faibles.

Si tous les indicateurs sont mutuellement indépendants, l'information globale devient la somme des puissances discriminatives de tous les indicateurs choisis, c'est-à-dire

$$C(1, 2, \dots, k) = \sum_{i=1}^k CD_i(P) \quad (5.2)$$

et si entre eux existe une dépendance fonctionnelle-linéaire, l'information globale se réduit à la puissance discriminative du premier indicateur, c'est-à-dire à

$$C(1, 2, \dots, k) = CD_1(P) \quad (5.3)$$

Si l'ensemble des indicateurs choisis peut être partagé en deux groupes mutuellement indépendants, l'information globale sera la somme des informations globales correspondant à ces deux groupes d'indicateurs.

Il faut encore souligner que, d'après l'expression analytique de l'information globale (5.1), les puissances discriminatives de tous les indicateurs représentent des quantités additives et que l'utilisation du facteur de pondération

$$(1 - r_{1i}) (1 - r_{2i.1}) (1 - r_{3i.12}) \dots (1 - r_{i-1,i.12\dots i-2})$$

fait soustraire la quantité d'information de l'indicateur X_i déjà contenue dans l'information des indicateurs qui lui précèdent par ordre de leur importance.

6. – LISTE OPTIMALE D'UN NOMBRE REDUIT DES INDICATEURS DE DEVELOPPEMENT.

La question fondamentale qui s'impose dans le problème du choix d'une liste des indicateurs de développement est la suivante : Si nous disposons d'une liste maximale de n indicateurs, comment la réduire en une liste de k indicateurs ($k < n$), mais de telle façon que l'ensemble de ces k indicateurs contienne une quantité d'information globale sur le niveau de développement des pays observés plus élevée que n'importe quel autre ensemble de k indicateurs choisis parmi les n indicateurs de la liste maximale.

Supposons que les indicateurs $\{X_1, X_2, \dots, X_n\}$ soient déjà rangés d'après le classement définitif et que

$$\langle \{X_{i_1}, X_{i_2}, \dots, X_{i_k}\} \rangle$$

est un sous-ensemble ordonné de k indicateurs, où $i_1 \leq i_2 \leq \dots \leq i_k$ et

$$r(X_{i_1}; D) \geq r(X_{i_2}; D) \geq r(X_{i_k}; D) .$$

La contribution globale de cet ensemble d'indicateurs dans l'évaluation du niveau de développement des pays est, d'après (5.1),

$$C(i_1, i_2, \dots, i_k) = \sum_{s=1}^k CD_{i_s}(P) \prod_{j=1}^{s-1} (1 - r_{i_j i_s \cdot i_1 i_2 \dots i_{j-1}}), \quad (6.1)$$

La liste optimale de k indicateurs sera constituée alors par le groupe $\{X_\alpha, X_\beta, \dots, X_\kappa\}$ qui maximise l'expression (6.1). En d'autres termes, la liste optimale sera formée par les indicateurs

$$X_\alpha, X_\beta, \dots, X_\kappa,$$

si

$$C(\alpha, \beta, \dots, \kappa) = \text{Max}_{i_1, i_2, \dots, i_k} \{C(i_1, i_2, \dots, i_k)\} \quad (6.2)$$

Même en utilisant un ordinateur, la détermination de cette liste optimale ne sera pas facile si la liste maximale contient un grand nombre d'indicateurs.

Cependant, dans la pratique, de hautes instances —telles que, par exemple, l'Assemblée générale des Nations Unies, ou le Conseil pour le Commerce et le Développement de la CNUCED ou l'ECOSOC, etc.— qui, en définitive, décident et adoptent une liste d'indicateurs de développement, pourraient imposer d'avance un certain nombre d'indications et laisser aux experts le soin de compléter cette liste afin de la rendre optimale. Peut-être que cette liste ne serait plus la meilleure possible, mais sur le plan technique, il serait bien plus facile d'exécuter des calculs.

Soit X_1, X_2, \dots, X_p les p indicateurs fixés d'avance et $X_{i_{p+1}}, X_{i_{p+2}}, \dots, X_{i_k}$ les $(k-p)$ indicateurs qu'il faut choisir parmi les $(n-p)$ indicateurs de la liste maximale. Pour une combinaison de $(k-p)$ indicateurs, nous pouvons ordonner l'ensemble formé par l'union de cette combinaison et les p indicateurs fixés d'avance, c'est-à-dire

$$\langle \{1, 2, \dots, p\} \cup \{i_{p+1}, i_{p+2}, \dots, i_k\} \rangle.$$

L'ordre de ces indices sera établi d'après le degré d'importance des indicateurs correspondants.

La liste optimale sera maintenant constituée par les p indicateurs fixés d'avance et les $(k-p)$ indicateurs $\{X_\nu, X_\mu, \dots, X_k\}$ pour lesquels

$$C[\langle \{1, 2, \dots, p\} \cup \{\nu, \mu, \dots, k\} \rangle] =$$

$$\text{Max}_{i_{p+1}, \dots, i_k} C[\langle \{1, 2, \dots, p\} \cup \{i_{p+1}, i_{p+2}, \dots, i_k\} \rangle] \quad (6.3)$$

Enfin, si les k indicateurs sont déjà choisis, soit en les fixant d'avance soit en les identifiant par le procédé (6.2) ou par le procédé (6.3), on pourrait se poser la question de savoir comment identifier l'indicateur X_{k+1} , qui apporterait l'accroissement maximum à l'information globale $C(1, 2, \dots, k)$.

Ce sera l'indicateur X_ω si

$$C[\langle \{1, 2, \dots, k\} \cup \{\omega\} \rangle] = \text{Max}_{i_{k+1}} C[\langle \{1, 2, \dots, k\} \cup \{i_{k+1}\} \rangle] \quad (6.4)$$

pour $i_{k+1} \in \{k+1, k+2, \dots, n\}$.

Si l'on considère séparément le P.I.B. par habitant, le groupe d'indicateurs économiques et le groupe d'indicateurs sociaux, la liste optimale contiendrait le P.I.B. par habitant, le groupe $\{X_\alpha, \dots, X_\kappa\}$ et le groupe $\{Y_a, \dots, Y_k\}$ pour lesquels

$$C(0; \alpha, \dots, \kappa; a, \dots, k) = \underset{i_1, \dots, i_\kappa}{\text{Max}} \underset{j_1, \dots, j_k}{\text{Max}} \{C(0; i_1, \dots, i_\kappa; j_1, \dots, j_k)\} \quad (6.5)$$

Remarquons, cependant, qu'il faudrait, dans ce cas-ci, calculer le coefficient de corrélation collective R_{xy} pour chaque combinaison $\{X_{i_1}, \dots, X_{i_\kappa}\}$ et $\{Y_{j_1}, \dots, Y_{j_k}\}$.

7. — CONCLUSION

Dans cette étude, nous avons essayé de présenter un procédé permettant d'établir une liste optimale d'indicateurs du développement dont les statistiques serviraient pour l'évaluation et les prévisions du développement dans la Deuxième Décennie.

On commence par l'établissement de la liste maximale des indicateurs de développement en réunissant les différentes listes utilisées jusqu'à présent ainsi que d'autres indicateurs qui ont été négligés ou ignorés dans le passé.

Le procédé consiste, ensuite, à évaluer quantitativement la puissance discriminative ainsi que l'importance de chaque indicateur de cette liste maximale.

On définit l'information globale d'un groupe d'indicateurs comme une somme de coefficients de la puissance discriminative de tous les indicateurs de ce groupe dont les coefficients de pondération correspondent à l'importance des indicateurs et éliminent les redoublements des mêmes quantités d'information. La liste optimale d'un nombre réduit d'indicateurs de développement sera alors formée par les indicateurs qui maximisent la quantité de l'information globale.

Etant donné l'absence ou la mauvaise qualité des données statistiques d'un grand nombre d'indicateurs de la liste maximale, le coût et le temps nécessaires pour introduire de nouveaux indicateurs dans le système statistique des pays en voie de développement, ce procédé permet de faire un plan rationnel dans la détermination de la liste d'un nombre limité d'indicateurs dont l'information globale est optimale.