JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

J. DUFRÉNOY

Introduction au vocabulaire de la statistique appliquée

Journal de la société statistique de Paris, tome 113 (1972), p. 261-272
<http://www.numdam.org/item?id=JSFS_1972__113__261_0>

III

VARIÉTÉS

INTRODUCTION AU VOCABULAIRE DE LA STATISTIQUE APPLIQUÉE

1. *La puissance des mots*

Le Reader's Digest consacre deux pages de chaque numéro à une rubrique « Cela paye d'enrichir votre vocabulaire » en invitant le lecteur à reconnaître le sens de chacun de vingt mots clés, utilisés dans le texte d'un article récemment publié : le 20e mot du numéro d'août 1971, page 141 est « contingent » qui s'écrit et se prononce de même en anglais et en français, et signifie « dépendant pour sa réalisation de quelque chose qui peut ou non se produire ».

Le même numéro annonce la publication, sous le titre « How to Increase your Word Power » (comment enrichir votre vocabulaire) d'un livre en deux volumes; le volume I indiquant les « racines » des mots (c'est-à-dire les matériaux utilisés pour construire chaque mot) et les règles d'utilisation correcte d'un mot scientifique, technique, ou même d'argot; le volume II constituant un lexoguide de 10 000 « mots clés ».

Enrichir son vocabulaire permet d'utiliser le mot propre dans chaque situation et, pour chaque message, parlé ou écrit, augmente les chances de succès, diminue les risques d'incompréhension ou d'interprétation erronée, et favorise les contacts efficaces (« bridge the communication gap »). Les résultats de divers tests, suggèrent une forte corrélation entre « maîtrise du vocabulaire » (vocabulary) et « rang social » (social standing), « Q. I. », « salaire ou revenu »...

L'américain moyen, ayant dépassé l'âge de 25 ans, n'acquiert guère plus de vingt-cinq mots par an alors que submergé, par la presse quotidienne et les « magazines », il est exposé, chaque jour, à plus de 10 000 mots imprimés, auxquels s'ajoutent la dizaine de milliers de mots qu'il entend chaque jour, en écoutant la radio pendant 75 minutes et en regardant la T. V. pendant des heures, ce qui implique la sollicitation d'au moins 560 réclames commerciales, relatives notamment à l'un des 200 détergents, des 350 aliments congelés, des 200 « farines à gâteau »...

Chaque année les progrès de la science et de la technique exigent l'introduction dans le vocabulaire d'un nombre de plus en plus grand de mots nouveaux dont la signification est d'autant plus accessible que l'on dispose déjà d'un vocabulaire plus étendu.

2. Le mot *contingent* est l'un des nombreux mots-clés qui ont même signification dans plusieurs langues; les mots clés d'usage international fréquemment employés dans le vocabulaire de la statistique peuvent être groupés quant aux notions que chacun permet d'exprimer; à chacun, est comparé un autre mot clé ayant à peu près la même signification (*cf.*) ou une signification opposée (VS).

Un mot clé contient et transmet de l'Information; en statistique cette information s'exprime par des valeurs numériques (résultats de dénombrements (fréquences f ou n) ou de mesures (y)) et permet « d'estimer avec précision le degré d'imprécision de nos connaissances ».

3. *Différences et proportions*

Nos connaissances sont « relatives » à des « différences » : dès 1755, Jean-Jacques Rousseau dans son « Discours sur l'inégalité » remarquait qu'« il y a loin du sentiment que tout être vivant peut avoir d'une différence, jusqu'à la conscience expresse d'une *relation* ou d'un rapport » (cité par M. Grot, État de nature, raison, progrès, selon le « Discours sur l'inégalité » et la « Profession de foi du vicaire... » *Revue de Synthèse* XC, janvier-juin 1969, pp. 5-30).

Une différence peut correspondre à l'alternative « tout ou rien », « oui ou non », « vivant ou mort », « débit ou crédit » (qui fournit a, qui reçoit doit); le « passif » peut s'exprimer par un chiffre affecté du signe « moins », l'« actif » par un chiffre affecté du signe « plus », la ligne de partage correspond à zéro, au-dessus de quoi se situent les « plus » au-dessous les « moins », dans un « système de référence » à une dimension; les chiffres affectés du signe plus (+) ou du signe moins (—) s'utilisent conformément aux règles de l'algèbre; la règle des signes, peut facilement s'exprimer par un graphique :

1° Au bas d'une feuille de papier, traçons une droite horizontale sur laquelle nous portons, à partir de l'origine (zéro à gauche) et vers la droite, les valeurs d'une variable X, en abscisses; vers le bord gauche de la feuille, traçons une droite verticale, le long de laquelle nous portons à partir de l'origine zéro, et vers le haut, les valeurs Y d'une deuxième variable.

Si, à chaque valeur x_i donnée à la variable indépendante X, correspond une valeur y_i de la variable dépendante Y, il existe une relation fonctionnelle entre Y et X_i; la relation la plus simple étant celle de proportionnalité $Y = bX$

2° Au lieu de ne considérer que la possibilité pour X de prendre à partir de l'origine zéro et vers la droite, des valeurs *positives* croissantes, et pour Y de prendre à partir de l'origine zéro et vers le haut, des valeurs *positives* croissantes, nous pouvons tracer la droite horizontale (axe des X ou abscisses) à mi-hauteur sur la feuille, et la droite verticale au milieu de la feuille, à égale distance des bords gauche et droit.

Nous divisons ainsi la feuille en 4 « quadrats » permettant d'illustrer la « règle des signes » de l'algèbre classique.

La verticale élevée du point $X = 0$ sépare le domaine des valeurs négatives de X (soit $X < 0$) du domaine des valeurs positives ($X > 0$); l'horizontale passant par le niveau $Y = 0$ sépare le domaine des valeurs négatives de Y (soit $Y < 0$) du domaine des valeurs positives ($Y > 0$) : si on adopte zéro comme limite supérieure de classe on distingue la classe des valeurs plus petites ou égales à zéro, c'est-à-dire atteignant au plus zéro ($\leqslant 0$) des valeurs supérieures (> 0); si on adopte zéro comme limite inférieure on distingue les valeurs situées au-dessous de zéro (< 0 de celles qui sont supérieures ou égales à zéro ($\geqslant 0$).

Valeur du seuil

Au lieu de zéro on peut utiliser une valeur critique (C) ou « seuil » telle que toute valeur de Y supérieure ou égale soit $Y \geqslant C$ soit considérée comme « normale » toute valeur inférieure $Y < C$ comme déficitaire ou anormale.

La survie d'un être vivant dépend du fonctionnement de chacun de divers systèmes « enzymatiques »; chaque système enzymatique est caractérisé par son aptitude à accomplir,

par unité de temps, une certaine quantité (y) de transformation chimique; chez l'individu normal (y) convenablement mesuré, doit dépasser un certain seuil (c); si y est nettement déficitaire par rapport à c, l'individu peut :

a) manifester des symptômes caractéristiques de cette déficience, ou

b) sans manifester de symptômes, être « porteur de la tare » et risquer de transmettre à un descendant la déficience, pouvant se manifester par des symptômes, ou conférant l'aptitude « porteur de tare ».

Risque évalué a priori

Dans une population humaine le risque d'être porteur d'une certaine tare peut être extrêmement faible (de l'ordre de 1 sur un million) mais pour une femme ayant mis au monde un enfant manifestant les symptômes de la tare, le risque de mettre au monde un autre enfant affecté des mêmes symptômes devient 1 sur 4; en prélevant dès les premiers mois de la grossesse un échantillon des tissus du fœtus pour mesurer l'activité (y) du système enzymatique suspect d'être déficient, il est possible de prédire que la grossesse aboutira à la naissance d'un nouveau né normal ($y > c$) ou risquera de produire un individu affecté des symptômes (y très inférieur à c).

Statistiquement, si $y \geqslant c$ on forme l'hypothèse nulle : L'activité enzymatique du fœtus ne diffère pas de la normale; si $y < c$ l'hypothèse nulle devient : le fœtus appartient à la catégorie des déficients risquant de manifester les symptômes de la tare, et il est à conseiller d'interrompre la grossesse; on prend le risque de faire l'erreur de 1re espèce, adopter comme vraie l'hypothèse nulle, quand elle est fausse, en utilisant l'information *a priori* obtenue par la mesure de y :

a) la mère décide d'interrompre la grossesse; le fœtus extrait chirurgicalement est soumis aux examens de laboratoire; le dépistage des symptômes attribuables à la déficience de l'enzyme fournit l'information *a posteriori* justifiant le choix de l'hypothèse acceptée *a priori* comme « vraie »;

b) on laisse la grossesse se poursuivre : la manifestation des symptômes imputables à la déficience enzymatique apporte *a posteriori* une justification de l'hypothèse adoptée comme *a priori*.

Ces exemples montrent comment les progrès de la technique, permettent d'une part d'augmenter l'information *a priori* quant au choix de l'hypothèse présumée « vraie », et d'autre part de déterminer *a posteriori* si ce choix a comporté ou non une erreur de première espèce.

Cette justification *a posteriori* du choix de l'hypothèse vraie est rarement possible; le statisticien doit le plus souvent se contenter de choisir pour représenter le phénomène étudié, le modèle mathématique le plus « vraisemblable » et la plus « simple » (celui qui comporte le plus petit nombre de paramètres). Nous prendrons comme exemple les Modèles de croissance économique ou de croissance du revenu national.

Théories de la croissance dans différents systèmes sociaux

Dans un certain système le revenu national Y a connu du début à la fin de l'année une augmentation $\Delta Y = \frac{1}{m} I - aY + uY$, selon un modèle mathématique n'impliquant que 3 coefficients (ou paramètres) : m = (capital) / (output); le niveau I d'investissement (avant déduction de la dépréciation) permet une production $\left(\frac{1}{m} I\right)$; la fiabilité (ferraillage

des équipements) diminue la productivité de $-aY$, tandis qu'une amélioration des possibilités existantes permet un accroissement uY du revenu national dont le taux de croissance

$$r = \frac{\Delta Y}{Y} = \frac{1}{m}\frac{I}{Y} - a + u$$

1) Économie socialiste : m, a et u sont déterminés relativement aux ressources; m et a dépendent du choix administratif quant au capital I_t à investir pour une nouvelle production (i) et quant au ferraillage; l'organisation scientifique du travail augmente la productivité (u) de l'équipement.

2) Économie capitaliste : selon la politique du laisser-faire, l'utilisation (u) de l'équipement varie selon la demande et peut augmenter ou diminuer au cours de cycles et le rapport m peut, lui-même, être affecté par la loi de l'offre et de la demande.

Le modèle ci-dessus ne fait pas apparaître ce qui, dans les ressources *(input)* utilisées pour obtenir l'*output*, est imputable à l'équipement matériel (à la technologie) ou à la main-d'œuvre.

Si la productivité de la main-d'œuvre est évaluée en « quantité de biens commercialisables produite par heure », toute augmentation de salaire horaire provoque une diminution de $\left(\frac{1}{m} I\right)$ et donc l'inflation, contre quoi une économie dirigée peut s'efforcer de lutter en bloquant les prix à la consommation (ce qui diminue la valeur monétaire de l'output) en en bloquant à la fois les salaires et les prix, ce qui stabilise (m).

Ordre, Information, Néguentropie

Le Monde est une immense accumulation d'ordres très divers dont le caractère commun est d'être improbable; nous nous refusons à admettre que tous les Ordres n'aient pas leur raison d'être, même si celle-ci devait être la raison non raisonnée, qui a nom Hasard (Jacques Rueff, Les dieux et les rois, pp. 101-102). Devant l'immense dose d'ordre, si immensément improbable que représente l'Univers, l'intelligence humaine a exigé une explication : elle en a obtenu plusieurs, relativement à la thèse de la Création et quant à la conception scientifique du Hasard et de la Sélection naturelle ».

L'intelligence humaine peut s'appliquer à interpréter l'ordre universel comme soumis à des lois naturelles qui peuvent chacune se représenter par des modèles mathématiques dont la vraisemblance s'estime par l'analyse statistique.

Au contraire, à des fins démagogiques, l'activité humaine peut s'opposer au fonctionnement des lois naturelles, viser à bouleverser l'ordre naturel, à maximiser le désordre, ce qui, selon la théorie de l'information, correspond à maximiser l'entropie, ou à minimiser la néguentropie, qui est une mesure de l'information.

Pour un système donné, le maximum d'incertitude, c'est-à-dire le minimum de néguentropie ou d'information, correspond à l'équiprobabilité de la chance de succès ou du risque d'insuccès : le jeu de pile ou face, avec à chaque coup 50 % de chances de pile et 50 % de face, peut être poursuivi mille fois sans que soit augmentée l'information relative à l'issue du mille et unième coup.

1. A San Francisco, en juin 1971, à la fin de l'année scolaire, 80 % des écoliers noirs fréquentaient 27 des 97 écoles élémentaires, le pourcentage des écoliers noirs dans ces 27 écoles variant de 47,3 à 96,8 % : dans le but de réaliser un mélange approchant de 50 % noirs et 50 % blancs, les 46 000 écoliers inscrits dans les 97 écoles ont été, par programmation sur calculatrice, répartis de telle sorte que plus de la moitié d'entre eux, au lieu de continuer

à fréquenter l'école la plus proche de leur domicile, seront transportés en autobus pendant parfois une demi-heure, vers une école éloignée, afin d'y réaliser le mélange 50 %, c'est-à-dire le maximum d'entropie, ce pourquoi le budget des écoles sera grevé d'une dépense supplémentaire de 1,6 millions de dollars, à récupérer par augmentation des taxes locales, à quoi s'ajouteront pour les parents voulant éviter à leurs enfants d'être « déportés » les frais d'inscription dans une école privée.

2. Le triomphe ou le défi de la statistique (J. L. Kahn *Rev. des Deux Mondes*, août 1970, p. 480) « L'instauration du suffrage universel fit franchir à la démocratie un grand pas; les lois des grands nombres donnaient des résultats qu'on ne pouvait prévoir; les enquêtes auprès de l'opinion publique ont fait franchir un nouveau pas : le choix d'un programme politique devient un jeu statistique; dans le nuage qu'on peut, dès lors appréhender, des opinions, il faut à tout moment, déceler les densités, les courants... Apparemment le triomphe de la statistique est le triomphe de la démocratie, mais... le jeu à ses périls ».

Un événement apporte d'autant plus d'informations qu'il est plus rare

L'une des milliers de boîtes de soupes fabriquées le 21 mai 1971 par la Compagnie « Bon Vivant » à Newmark et étiquetées « Vichyssoise » a causé le 30 juin un mort à New York; une intoxication mortelle du botulisme (la contamination par le microbe du botulisme) résultant de stérilisation insuffisante; quelques semaines plus tard la Compagnie faisait faillite. Quelques mois plus tard, les services de Recherches de la Campbell Soup Co décelaient une contamination par microbe du botulisme dans un lot de 12 000 boîtes de soupe « poulet avec légumes » (chicken vegetable soup) fabriquées le 15 juin dans l'usine de Paris, Texas; l'alerte fut aussitôt donnée aux Services fédéraux du ministère de l'Agriculture (USDA) (dont un agent doit personnellement surveiller la fabrication de telles conserves) ainsi que la *Federal Drug Administration* (FDA) en vue de prévenir la consommation de la soupe du lot incriminé (n° de code 07. P13,701 X) dont 54 % des boîtes ont déjà été récupérées; la Compagnie rembourse les boîtes qui lui seront retournées. Aucune intoxication n'a d'ailleurs été signalée et les actions de la Compagnie n'ont subi qu'une baisse insignifiante. La stérilisation correcte d'une boîte de conserve, compte tenu des risques de contamination par « microbe résistant à la chaleur » implique le séjour pendant (m) minutes dans une enceinte chauffée à (t^0); au-dessus d'un certain seuil (mt) la stérilisation est considérée comme acquise, mais les qualités gustatives tendent à se dégrader à mesure que la conserve est soumise plus longtemps à température plus élevée, d'où intérêt de maintenir (mt) au minimum compatible avec l'élimination du risque, sauf accident de fabrication.

A l'industrie des conserves s'applique la loi des probabilités extrêmes : le risque pour une boîte d'être contaminé est tellement voisin de zéro qu'il peut être considéré comme nul; cependant les deux exemples précédents montrent qu'après une série de centaines de millions de boîtes correctes, une usine peut produire un lot contaminé ou en d'autres termes une usine ayant produit pendant une longue série d'années, des lots corrects, peut laisser sortir un lot contaminé : la probabilité (c) de zéro, un, deux... accidents de fabrication peut s'estimer conformément aux séries de Poisson : dans un intervalle de temps déterminé (t) *le plus probable est qu'il ne se passe rien*, soit $c = 0$; mais si l'on envisage un nombre suffisamment grand d'intervalles tels que (t) on évoque l'éventualité de $c = 1$, puis $c = 2$ manifestations d'accidents.

Le « contrôle » de la fabrication » utilise les distributions de séries de Poisson, notamment pour les études de « fiabilité » ou du « vieillissement ». Les progrès de la technique

permettent d'augmenter l'intervalle de temps (t) pendant lequel le plus probable est qu'il ne se passe rien ($c = 0$).

Les épreuves de fiabilité permettent de dépister les individus défectueux, dès l'origine *(dead on arrival)* à éliminer pour ne soumettre aux épreuves de vieillissement que les individus « capables de vieillir » : en microélectronique d'après Geoffroy Dunner (New scientist and science Jour. 8 july 1971, p. 75) 5 % des pièces soumises au contrôle sont défectueuses à l'origine (dead on arrival); parmi les autres le taux de mise hors service (failure rate) est 0,01 % par mille heures de fonctionnement : on peut donc prédire une déficience (ou panne) en 40 ans pour un système comportant 300 circuits.

La croissance économique atteint dans certains pays une vitesse telle que l'enfant de dix ans vit dans un « environnement » comportant deux fois plus de produits manufacturés qu'il n'en existait à sa naissance; l'intervalle de temps entre « découverte scientifique », « réalisation technologique » et « commercialisation de la fabrication en série » est passé de 34 ans pour les aspirateurs, les fourneaux électriques, les réfrigérateurs d'avant 1920, à 8 ans pour les appareils de télévision et les « machines à laver et à sécher » (washer-dryer) commercialisés de 1939 à 1959; maintenant cet intervalle se réduit à quelques mois pour les transistors et divers dispositifs électroniques.

A mesure que deviennent disponibles des biens de plus en plus « durables » la société s'oriente vers la consommation d'objets ne servant qu'une fois (emballages perdus, serviettes en papier, repas préfabriqués présentés dans une « vaisselle » qu'on jette après le repas (TV dinners are served in throw-away trays).

Chacun doit avoir dans la tête un modèle mental du monde : pour pouvoir « fonctionner » ou même « survivre » il faut adopter un modèle qui jouisse de quelque vraisemblance, c'est-à-dire qui fournisse une représentation adéquate du présent et permette de « prévoir »; dans notre société actuelle où la « vérité d'hier » devient « fiction » aujourd'hui et vice versa, beaucoup sont tentés d'adhérer à telle ou telle idéologie, imaginée par tel ou tel individu ignorant qu'il existe une technique, une méthode, permettant d'évaluer le « degré de vraisemblance » d'une méthode imaginaire; certains cependant s'astreignent à passer par la « Porte étroite » qui par l'analyse statistique permet de parvenir à « estimer avec précision le degré d'imprécision de nos connaissances » à calculer le risque inhérent à toute décision ou à tout choix en face d'une alternative et à prédire, dans l'intervalle des prévisions optimiste, ou pessimiste, l'évolution la plus probable d'un phénomène et d'élaborer logiquement une « stratégie ».

Annexe

Sous le vocable d'« intégration » et dans le but de rendre dans chacune des 97 écoles de San Francisco, aussi semblables que possible les pourcentages d'écoliers de chaque groupe ethnique (W, blancs, N, Noirs, C, Chinois, AL, Amérique Latine) on « déporte » en autobus des enfants d'un groupe vers des écoles où prédominent les enfants d'autres groupes ethniques; un sondage a été effectué dans un échantillon de parents, 35,3 % W, 30,7 % N, 17 % C et 17 % AL : les questions posées étaient :

1. Considérez-vous l'éducation que reçoivent vos enfants, comme excellente (E), bonne (B), satisfaisante (S) ou mauvaise?

2. Approuvez-vous ou non le programme de transport par autobus?

3. Nos enfants bénéficieront de l'expérience acquise dans une école racialement équilibrée?

4. Le transfert par autobus privera nos enfants de la participation aux activités parascolaires.

5. Je suis persuadé que mon enfant sera bien traité et compris par les enseignants d'une école éloignée de notre quartier;

6. Le transfert de mon enfant vers une école éloignée me fera perdre le contact avec l'école.

7. Je considère que l'école que fréquente mon enfant reflète :

a) les caractéristiques raciales de notre quartier,
b) les caractéristiques économiques du quartier;

8. Je suis pour l'intégration mais contre le transfert par autobus;

9. J'aurais des inquiétudes quant à la sécurité de l'enfant :

a) transféré à une école lointaine,
b) transporté chaque jour en autobus;

10. Je considère que mon enfant recevra une meilleure éducation dans une école autre que celle de notre quartier.

11. Je préfère une école non intégrée.

Pour chacune des questions les pourcentages de oui et de non sont indiqués pour chaque groupe ethnique

Question	2		3		4		5		6		7a		8		9a		10	
	Oui	Non	Oui	Non	Oui	Non	Oui	Non	Oui	Non	Oui	Non	Oui	Non	Oui	Non	Oui	Non
W	14	83	33	51	75	17	59	28	88	11	60	33	88	10	75	23	13	73
N	39	56	64	24	51	39	67	18	53	44	37	50	50	37	50	45	46	43
LA	38	59	59	21	46	44	60	23	71	23	50	36	73	23	88	10	29	54
C	6	92	28	54	79	17	23	36	82	6	58	27	96	4	92	6	2	67

Pour chacune de ces questions (alternative oui ou non) les pourcentages de « sans opinion » sont relativement faibles.

La question 12 « Pensez-vous que les écoles intégrées provoqueront trop de contacts entre races? les pourcentages de réponse sont :

	oui	non	sans opinion
W	15	75	8
N	7	80	12
C et LA	33	52	

La question 13 « Pensez-vous que réaliser l'équilibre racial par transfert d'écoliers en autobus :
— élèvera le standard d'éducation beaucoup ou un peu;
— l'abaissera beaucoup ou un peu;
— n'aura pas d'effet.

	élèvera	abaissera	sans effet
W	7	43	44
N	32	12	40
LA	30	31	35
C	2	44	35

Information et culture

De nos jours, ce qui manque aux cadres, aux techniciens, aux chercheurs, ce ne sont pas les connaissances intrinsèques, ni les idées, mais, le plus souvent, l'aptitude à ordonner ces connais-

sances, à mettre ces idées en perspective, à exprimer les unes et les autres, noir sur blanc, en français (ou en anglais) correct, clair, intelligible.

Tanguy Kenec'hdu, Anglais scientifique ou Science de l'anglais? (*Rev. des Deux Mondes*, août 1970, pp. 391-394).

L'information est un mode d'acquisition de savoir, de connaissances... la culture ne commence qu'au-delà, à un niveau d'assimilation, d'interprétation, d'utilisation du savoir et des connaissances... ce qui fait la différence entre l'esprit humain et le cerveau électronique, entre l'homme et l'ordinateur (*Rev. des Deux Mondes*, août 1970, p. 492).

Le dessin est un moyen de noter, de comparer les multiples informations nécessaires à l'exercice d'une activité moderne, mais peu de personnes savent *utiliser* le dessin; un individu « scolarisé » a consacré quelques 500 heures à *reproduire* un objet, il n'a pas consacré un instant à dessiner en vue de prévoir la transformation des individus et des choses, par expression graphique, selon les techniques de la Sémiologie graphique (diagrammes, réseaux, cartes...), décrits par Jacques Bertin de l'École pratique des hautes études (Labo. de Cartographie dans un livre publié chez Gauthiers-Villars, 1967).

Explication, prédiction (prospective) et théorie de l'information

Une prédiction scientifique comporte :

1° Une description des conditions dans l'état présent, avant la prédiction.

2° L'utilisation d'une loi générale « déterministe » ou « stochastique » (faisant intervenir des considération statistiques).

Exemple : Un stimulus (S) provoquera une réaction (r) chez un certain individu; la situation initiale est définie par le stimulus (S), la prédiction est relative à la réaction (r), qui dans le cas le plus simple peut s'exprimer dans le code binaire par 0 ou 1 : l'un de deux événements ROUGE (R) ou VERT (V) peut se manifester au sujet de l'allumage de l'une de deux lampes l'une rouge, l'autre verte; l'individu doit prédire en appuyant sur l'un des deux boutons marqués R ou V lequel des deux événements se produira au cours de l'expérience, conçue par l'expérimentateur selon une séquence prédéterminée d'allumages de rouge ou de vert; chaque séquence possible Rouge, Rouge, Vert Rouge, Rouge Vert ou Vert vert, représente un stimulus pour chacun des n essais; au bout de n essais, la séquence de n allumages « rouge » ou « vert » correspond au stimulus total, c'est-à-dire aux conditions initiales de la prédiction au n_i^{eme} essai; soient S^* une séquence spécifique de réactions (chacune exprimée en poussant un bouton) on fait cette prédiction : avec une probabilité p, le stimulus S^* produira la réaction r^*, ou $P\,(r^*/S^*) = p$.

Une observation, relative au résultat d'une certaine expérience, contient une quantité d'information qui dépend en partie de ce que sait et croit celui qui a fait l'observation, et qui dépend notamment du degré d'imprévu ou d'« invraisemblance. »

Comme mesure numérique de l'information, Shannon a proposé le logarithme de la vraisemblance; exemple : une expérience, effectuée dans les conditions initiales définies par (S) devrait produire un résultat (r); la probabilité *a priori* d'obtenir le résultat (r) sous l'effet du stimulus (s) étant $P_0\,(r \mid s)$, le contenu *a priori* d'information de l'expérience (e) pour une certaine réaction (r) est

$$I_0 = \log_K \left[\frac{1}{P_0\,(r \mid s)}\right] \qquad \text{lorsque } K \text{ est l'unité d'information.}$$

On en déduit le coefficient de puissance de prédiction.

$$I_\alpha = I_0 - \log_K \left[\frac{1}{P_\alpha (r \mid s)}\right]$$

lorsque $P_\alpha(r \mid s)$ est la probabilité *a posteriori*, estimée d'après un modèle mathématique M.

En choisissant comme unité d'information l'inverse de la vraisemblance *a priori*, c'est-à-dire en écrivant

$$k = \frac{1}{P_0\,(r \mid s)} \qquad \text{et en écrivant} \qquad I_\alpha = 1 - \log_{KL} \frac{1}{P_\alpha\,(r \mid s)}$$

on *normalise* la mesure de l'information : le coefficient de puissance de prédiction devient le pourcentage de l'information *a priori* dont rend compte le modèle.

Théorie de l'information et thermodynamique

Principe néguentropique de Brillouin : L'information dite liée apparaît en tant qu'opposée à l'entropie totale du système physique.

1er principe de thermodynamique, établi par Carnot :

« La quantité totale d'énergie d'un système ne change pas lorsque l'une des formes (chimique, mécanique, électrique, thermique) se change en une autre. »

2e principe, énoncé par Clausius en 1850 :

Dans tous les processus naturels (hormis ceux de la matière vivante) les conditions de température et de pression restant constantes, une certaine fraction de l'énergie libérée ne peut être utilisée pour l'accomplissement d'un travail : cette fraction impossible à utiliser « isothermiquement » (à température constante) dont la valeur peut être calculée en fonction du zéro absolu (— 273 °C) a reçu le nom d'entropie : un système constitué par un morceau de sucre dans de l'eau subit une augmentation d'entropie (une diminution de néguentropie) à mesure que le sucre se dissout dans l'eau. Seule la fraction « ordonnée » de l'énergie peut être transformée en travail : l'entropie mesure le degré de désordre.

Les êtres vivants tendent à diminuer l'entropie, et à augmenter l'ordre, c'est-à-dire la néguentropie.

Maxwell avait imaginé un « démon » posté à l'orifice minuscule, faisant communiquer deux enceintes A et B emplies de gaz à une température, et manœuvrant une trappe lui permettant d'interdire le passage d'une enceinte à l'autre de telle ou telle molécule. S'il laissait passer de A vers B les molécules « rapides » (à haute énergie) et de B vers A les molécules lentes (de faible énergie), il faisait monter la température de A, baisser celle de B; pour que le démon puisse fermer la trappe en connaissance de cause il doit avoir *mesuré* la vitesse de chaque molécule; chaque mesure, c'est-à-dire chaque acquisition d'information consomme de l'énergie; ce qui compense la diminution d'entropie du système, conformément à l'équivalence entre information et néguentropie.

Dans la cellule vivante chaque « enzyme » fonctionne à la manière du « Démon de Maxwell » pour créer de l'ordre au prix d'une consommation d'énergie potentielle chimique : la consommation d'énergie est d'ailleurs extrêmement faible, par comparaison avec les échanges d'énergie provoqués par le fonctionnement de l'enzyme, asservie elle-même à un « effectent » : l'activité de l'enzyme n'est pas proportionnelle à la concentration de son effectent; dans la phase initiale de la réaction l'effet de l'activateur croît beaucoup plus vite que sa concentration pour atteindre une valeur maximale et décroître ensuite selon une courbe en S (sigmoïde) : de même en électronique, il importe que la réponse d'un relai soit non linéaire par rapport aux variations de potentiel qui le gouvernent. Mais pour réaliser des performances analogues à celle d'un relais électronique de masse (M) il peut suffire d'une masse d'enzyme (M) (10^{-9}) dans les conditions actuelles de « miniaturisation » (d'après J. Monod : Le Hasard et la Nécessité : essai sur la philosophie naturelle de la biologie moderne, Éditions du Seuil, Paris; v. La Pensée, no 155, février 1971.

Séries chronologiques

Représentation du temps mathématique : la notion du temps correspond à la notion d'un « changement », d'un « mouvement », de la position « d'avant » ou passé vers la position après : une ligne (trajectoire) étant parcourue par un « mobile », à chaque position successive du mobile correspond un point de la ligne, et réciproquement à chaque point correspond une des positions successives : ces positions considérées, abstraction faite du mobile, seulement selon leur rang, forment les termes d'une succession qui constitue le temps : la succession étant considérée en dehors du mouvement, les positions successives, comptées successivement 1, 2, ..., i peuvent être projetées sur une dimension idéale, qui constitue l'échelle du temps, le long de laquelle se succèdent les instants successifs $t_1, t_2, \ldots, t_i$. La succession se représente par une suite de termes correspondant chacun à l'un des points de l'espace parcouru; chacun de ces points est défini par le « lieu » occupé sur la trajectoire, à un instant t_i du temps mathématique : le temps mathématique est l'une des deux dimensions nécessaires à la représentation objective du mouvement (l'un des deux paramètres requis pour la mesure du mouvement), l'autre dimension (l'autre paramètre) correspondant à l'espace.

Il y a presque 500 ans, Nicole Oresme avait imaginé de porter la succession des intervalles (l'échelle des temps) sur une droite horizontale au bas d'une feuille, le « lieu » occupé par le mobile

à chaque temps t_t étant « projeté » sur une droite verticale élevée à partir de l'origine ($t = 0$) à gauche de la feuille. Cette représentation du mouvement utilisant deux axes de coordonnées (l'axe des abscisses pour les intervalles de temps, l'axe des ordonnées pour les intervalles d'espace) a été utilisée par Descartes, d'où le nom de coordonnées cartésiennes :

Mouvements périodiques : le mobile, partant de l'origine ($d = 0$) de l'échelle des distances à l'origine du temps ($t = 0$) s'éloigne jusqu'à une distance maximale, puis revient à son point de départ : la mesure des distances parcourues à chaque instant (t_t) doit se faire à partir d'un point fixe, choisi de telle sorte que le mouvement puisse se représenter de façon aussi simple que possible quant à chaque distance parcourue dans chaque intervalle de temps; le temps (t) étant considéré comme la variable indépendante et la distance parcourue (d) comme la variable dépendante on choisira le système de référence qui permet d'exprimer la relation fonctionnelle entre (d) et (t) par une formule aussi simple que possible.

Par exemple, le déplacement de la Terre par rapport au Soleil et aux autres corps célestes se représente le plus simplement selon l'hypothèse que la Terre tourne sur elle même en 24 heures et en 365 jours autour du soleil pris comme point fixe. Dès 1377, Nicole Oresme, traduisant à l'intention de Charles V le traité d'Aristote « de Celo » et le commentant, adoptait comme vraisemblable l'hypothèse du mouvement diurne de la Terre devant un ciel fixe : cette traduction a fait récemment l'objet de rééditions et traductions (Nicole Oresme, Le Livre du Ciel et du Monde, edited by A. D. Menut et A. J. Denomy, translated with an introduction by A. D. Menut, Madison, Milwaukee and London, The University of Wisconsin Press, 1968).

Séries chronologiques appliquées à l'Économie

Une série chronologique comporte le classement, par unité de temps (seconde, minute, heure, jour, année, siècle, millénaire...) de résultats de dénombrements (n) ou de mesures (y); sur l'échelle des temps (t), à partir de l'origine $t = 0$ (correspondant au temps présent), on peut porter vers la droite, (vers l'avenir) les temps à venir t_1, t_2, t_t, ..., t_K, et vers la gauche, symétriquement, vers le passé, les temps t_{-1}, t_{-2}, ..., t_{-t}, ..., t_{-K}.

L'échelle peut-être « arithmétique : à chaque intervalle d'une seconde, minute, heure, année, siècle... correspond alors une même distance (Δt) sur la droite où l'on porte l'échelle.

Mais on peut obtenir une représentation du temps en utilisant une échelle « arc tangente ».

Pour représenter graphiquement une série chronologique, ayant construit sur une droite horizontale, au bas de la feuille, l'échelle des temps (t) on construit sur une droite verticale, à gauche de la feuille, l'échelle des valeurs dénombrées (n) ou mesurées (y); on porte chaque valeur dénombrée (n_{-t}) ou mesurée (y_{-t}) à son rang chronologique sur l'échelle des temps, pour le passé, (t_{-1}, t_{-2}, ..., t_{-t}, ...).

La distribution des points chacun de niveau (n_{-t}) ou (y_{-t}) au moment ($-t$) de l'observation, indique une « tendance » (trend) permettant de « prévoir » donc de « prédire » par « prospective », l'allure probable (ou vraisemblable) de l'évolution des résultats de dénombrements (M_t) ou de mesures (y_t) au temps t dans l'avenir.

L'utilisation des séries chronologiques, en remontant dans le passé, permet de faire de l'Histoire quantitative, ou plutôt de l'Histoire sérielle, « qui s'intéresse moins au fait individuel qu'à l'élément intégrable dans une série homogène, susceptible de porter les procédés mathématiques d'analyse des séries et d'être raccordée aux séries qu'utilisent couramment les autres sciences de l'homme ».

1) « Les deux secteurs de l'histoire qui furent, en premier, touchés par le simple souci de compter, par quoi tout progrès passe, sont ceux de la population et des prix. »

La mesure est entrée dans l'histoire par les prix : dès 1688, Gregory King fournit, en s'appuyant sur des sondages, le premier schéma d'une comptabilité nationale, par l'évaluation du produit national de l'Angleterre et du Pays de Galles; cette date 1688 marque pour P. Chaunis (L'Histoire sérielle, Bilan et Perspectives, Revue historique 494, av. juin 1970, pp. 297-320) le début de « l'Europe des Lumières » et celui de l'ère protostatistique, succédant à l'ère préstatistique, où les éléments bruts (mercuriales, pour les prix, registres paroissiaux pour la population) demeurés inexploités, peuvent servir à allonger dans le passé, une série chronologique :

Par exemple J. Dupaquier, M. Lachiver et J. Meuvret (Mercuriales du Pays de France et du Vexin Français, 1640-1792, Paris S. E. V. P. E. N., 1968) ont rassemblé les séries chronologiques de prix des céréales (blé, seigle, avoine, orge) sur des marchés situés en plein cœur de riches régions

agricoles (Gonesse en Pays de France, Magny et Marines, en Vexin français) ou sur la Seine et sur l'Oise (Meulan et surtout Pontoise, le plus important de tous).

Un graphique où les prix sont portés en ordonnées sur échelle arithmétique révèle, sous les fluctuations dues à des séquences d'années de tendance à la hausse et à des séquences d'années de tendance à la baisse, une tendance générale à la hausse, au cours des sept décennies 1729 à 1739, 1739-1749... 1779-1789, et enfin au cours des 3 années 1790-1972(les indications de prix manquant de 1793 à 1799).

En fonction de la variable indépendante X, prenant successivement, par rapport à la décennie 1759-1769, pour laquelle on pose $\bar{x} = 0$, les valeurs $x = -3$ pour 1729-1739, $x = -2$ pour 1739-1749, $x = -1$ pour 1749-1759, et symétriquement $x = 1$ pour 1769-1779, $x = 2$ pour 1779-1789 et $x = 3$ pour 1789-1792, la relation fonctionnelle entre prix moyen Y_x pendant chacune des 6 décennies ou pendant la période 1780-1792, peut s'écrire, après transformation des Y en $y' = \log_e y$ sous la forme de l'équation : $\hat{y}' = \bar{y}' + B(x - \bar{x}) = 5\,271 + 0{,}05\,(x - \bar{x})$ équation de la droite représentant, après transformation $y' = \log_e y$, la courbe exponentielle de l'augmentation des prix y de 1729 à 1792. Puisque $\bar{y}' = 5\,271 = \log_e 195$ une droite tracée sur le graphique, partant du niveau $y = 155$ pour 1735, passant par le niveau $\bar{y} = 195$ en 1765 et qui atteindrait le niveau 240 en 1795, rend compte de la tendance à augmentation au cours des deux derniers tiers du XVIII^e siècle.

Par contre, au XIX^e siècle, du moins après 1810, se manifeste une tendance générale à la baisse, malgré les fluctuations dont l'amplitude tend à décroître de 1810 à 1870; les révolutions de 1830 et 1848 marquent chacune l'aboutissement d'une décennie de hausse.

Le tirage au sort du 4 décembre 1971, des jeunes gens nés en 1952 pouvant être appelés à servir dans l'armée des États-Unis a été effectué par prélèvements successif dans chacun de deux tambours contenant chacun 366 capsules; les 366 capsules du 1^er tambour renferment chacune la date de l'un des 366 jours de l'année bisextile; les 366 capsules du 2^e tambour l'un des numéros d'ordre de priorité de 001 à 366 : par exemple, à la date 4 décembre extraite du 1^er tambour a correspondu le numéro 001 extrait du 2^e tambour, c'est-à-dire que le risque d'incorporation est maximal pour les jeunes gens nés le 4 décembre 1952; par contre à la date du 1^er novembre correspond le n° 366 comportant le minimum de risque. En fait le risque d'incorporation est presque nul pour les numéros supérieurs à 160 ou même à 150; cependant les jeunes gens dont la priorité est inférieure à 175 sont soumis à l'examen physique en vue d'incorporation éventuelle : le tableau indique les fréquences pour chaque mois de tirages de numéros inférieurs à 175 et de numéros égalant ou dépassant 175, ainsi que le risque d'incorporation pour les jeunes gens nés dans chacun des 12 mois de l'année 1952.

	< 175	⩾ 175	t	r
Janvier	8	23	31	0.258
Février	12	17	29	0.416
Mars	12	19	31	0,387
Avril.	16	14	30	0.503
Mai	20	11	31	0.645
Juin	15	15	30	0.500
Juillet	17	14	31	0.548
Août	15	16	31	0.484
Septembre	13	17	30	0.433
Octobre	19	12	31	0.613
Novembre	10	20	30	0.333
Décembre	18	13	31	0.580
	175	191	366	0.477

Les numéros d'ordre étant rangés par ordre chronologique depuis le 1er janvier (nº 207) jusqu'au 31 décembre (nº 322) on peut dénombrer : 40 sorties d'un numéro égal ou supérieur à 175 entre numéros inférieurs (c'est-à-dire 40 absences de « suite » de numéros supérieurs à 175) 18 sorties de 2 numéros supérieurs à 175 pour deux dates successives...

Si l'on définit par zéro, l'absence de suite (un numéro égal ou supérieur à 175 est précédé et suivi de numéros inférieurs) par 1 la sortie successive de 2 numéros, tel qu'un numéro égal ou supérieur à 175 est suivi d'un autre numéro égal ou supérieur à 175, par 2 la sortie successive de 3 numéros (tels que 175 ou au-dessus)... on peut dresser les tableaux suivants :

			c	f	F	fc
1	40	40	0	40	97	0
2	18	36	1	18	57	18
3	16	48	2	16	39	132
4	7	28	3	7	13	21
5	3	15	4	3	6	12
6	1	6	5	1	3	5
7	0		6	0	2	0
8	1	8	7	1	2	7
9	0		8	0	1	0
10	1	10	9	1	1	10
		191				105

J. Dufrenoy