

LES MODELES MATHÉMATIQUES APPLIQUES DANS LA GESTION DES BIBLIOTHÈQUES

Mme Amel KHEDDACHE KASDI
CERIST

Ingénieur en Recherche Opérationnelle - Chargée d'études
Laboratoire Recherche & Développement en IST

Tel. : (213) 02 91 62 04-09 - Fax : (213) 02 91 21 26

Amel@Tassili.cerist.dz

RESUME

La gestion scientifique des bibliothèques a toujours été un souci majeur pour son gestionnaire. De formation documentaliste, ce dernier s'est vu " maladroit " dans l'exploitation de la théorie mathématique et plus particulièrement de la statistique.

A travers cet article, nous nous proposons, dans un premier temps, de tracer un historique montrant l'évolution de l'utilisation des modèles mathématiques dans la gestion et la prévision dans les bibliothèques. Ensuite, nous définirons et exposerons le principe de chaque modèle d'analyse ; et enfin, nous clorons notre article par une étude comparative aboutissant à la sélection de l'une des méthodes comme " solution " dans la gestion de la bibliothèque.

1. INTRODUCTION

La gestion des bibliothèques est de nos jours un processus complexe. Ceci est dû, en partie, à l'introduction des nouvelles technologies et à l'abondance des publications ; aussi, nous savons que gérer implique une évaluation et une planification.

L'évaluation met à la disposition du responsable de la bibliothèque, l'état actuel de son institution et ainsi donc dévoile les lacunes du système de gestion de celle-ci ; alors que la planification est la tâche la plus importante dans la gestion. Il s'agit de faire des prévisions.

La pratique a fait que la gestion des bibliothèques repose essentiellement sur une bonne évaluation et des prévisions à court terme basées sur un besoin ressenti.

Cette façon de faire s'est vue donner des résultats si corrects que les gestionnaires des bibliothèques ont fini par opter pour cette méthode de travail.

Ceci nous amène à se demander pourquoi le gestionnaire n'a jamais innové en recourant aux techniques de prévisions classiques ¹ sachant que " un modèle permet d'interpréter un ensemble de phénomènes au moyen d'une structure dont il montre les principaux éléments et les relations qui existent entre ces éléments " ² ?

Nous savons que la formation des bibliothécaires ne permet pas la manipulation des modèles théoriques de prévision ; aussi, dans cet article, nous nous proposons de définir quelques méthodes des plus utilisées dans la prévision dans les bibliothèques à l'échelle internationale et de sélectionner celle susceptible d'être employée dans le cadre algérien à partir d'un programme informatique.

2. HISTORIQUE

A la suite de la découverte des méthodes de gestion quantitative à savoir les

¹ Ces méthodes seront définies par la suite dans ce document.

² Définition de MCQUAIL et WINDAHL : Communication models for study of mass communications, Londres, Longman, 1986.

méthodes de Recherche Opérationnelle ³, les institutions ont été enthousiastes pour leur application.

Cependant, les bibliothèques ont été les dernières institutions à adopter cette gestion, bien que Ralph SHAW préconisa en 1947 une gestion scientifique dans une bibliothèque.

En 1954, SHAFFER et ERNT ont développé quelques modèles mathématiques déterministes pour l'estimation et la prévision de la valeur moyenne de l'utilisation des livres.

Dans un cadre plus général, J. W. FORRESTER a tenté de développer des méthodes originales dans le management des sociétés et plus particulièrement des "industrielles" et ce en 1961 ⁴.

En 1962, ROTHKOPF travailla sur la gestion des bibliothèques et développa lui aussi des modèles mathématiques déterministes concernant l'utilisation des livres. Cette année là, DAWSON, ALDRIN & COULD ont également traité ce sujet et ont fini par développer des modèles mathématiques de gestion ⁵.

Le premier livre sur la gestion scientifique dans les bibliothèques apparut en 1966 par DOUGHERTY et HEINRITZ. Mais le mérite en revient à Morse en 1968, car il fut le premier à publier une étude compréhensive, utilisant la théorie des probabilités et des techniques modernes analytiques, dans la prévision de l'utilisation des bibliothèques. Cette étude comprend même des résultats et demeure jusqu'à présent l'une des plus répandues.

Ce n'est que dans les années 70 que les bibliothécaires ont commencé à signaler l'importance des planifications méthodologiques dans leurs écrits.

³ Nous rappellerons que la Recherche Opérationnelle a été utilisée la première fois dans la stratégie militaire à la deuxième guerre mondiale.

⁴ J. W. FORRESTER a continué son étude et obtint en 1969 et en 1971 d'autres modèles.

⁵ En 1967, JAIN aussi développa des modèles mathématiques pour l'estimation et la prévision de l'utilisation des livres.

En 1970, HAYES et BECKER considèrent que les prévisions dans les bibliothèques faisaient impliquer en grande partie le jugement des professionnels de ces bibliothèques et donc étaient pessimistes quant à l'utilisation seule des modèles mathématiques.

Par la suite, en 1972, MORSE développa ses modèles de MARKOV de base et SWANSON et BOOKSTEIN confirmèrent que l'application des modèles valides demeurait une solution ingénieuse. L'une des études méritant d'être citée, dans ces années là et plus précisément en 1976, est celle de M. A. DRAKE. En effet, il considéra la régression linéaire comme une technique de prévision à appliquer.

En 1978, MAKRIDAKIS et WHEELWRIGHT donnèrent deux versions des prévisions par l'approche des séries chronologiques, alors qu'en 1979, J. H. KANG présentait l'une des plus sophistiquées des études de prévision dans la littérature des bibliothèques, au cours de cette décennie. Il étudia les prévisions des demandes en utilisant plusieurs méthodes, y compris celles incluant l'aspect cyclique des données et fini par conclure que la régression était la plus adéquate pour ce cas d'étude.

Le début des années 80 a vu apparaître des études faisant appel à des techniques si complexes que ROWLEY et ROWLEY ont, en 1981, blâmé la Recherche Opérationnelle qui avait tendance à utiliser des modèles trop complexes pour la compréhension des bibliothécaires en général et des gestionnaires des bibliothèques en particulier.

En effet, en 1980, HODOWANCE utilisa la régression multiple pour l'estimation du modèle de prêt des documents dans les bibliothèques publiques. Et depuis, les études sur la gestion des bibliothèques tendent vers des modèles de régression multiples et des modèles de séries chronologiques.

Ceci a fait qu'elles sont peu nombreuses et surtout redoutées par les gestionnaires des bibliothèques.

3. LA GESTION DE QUELQUES FONCTIONS D'UNE BIBLIOTHEQUE A PARTIR DE MODELES MATHEMATIQUES

3.1. Les différents modèles

Dans cette partie de l'étude, nous allons énumérer et définir les modèles basés sur des méthodes mathématiques, les plus utilisés jusque là dans la gestion des bibliothèques.

Cependant, nous rappellerons que notre souci n'est pas de définir explicitement les méthodes d'analyse ⁶ mais plutôt de définir leur principe.

a) *Modèles basés sur les méthodes analytiques :*

Ces modèles ont été des plus utilisés car ils se basent sur de simples formules mathématiques susceptibles de répondre au besoin du gestionnaire.

Nous citerons à titre d'exemple :

- pour déterminer le coût moyen d'achat d'un document (noté coût) :
coût = (coût moyen d'un livre X % usagers) + (prix moyen d'une revue X % usagers x un complément d'universités et d'étudiants) ⁷
- pour déterminer le nombre total de documents achetés relativement à un budget alloué (noté nombre) :
nombre = ((le montant du budget total alloué +/- la variation dans ce budget) - (une constante coût ⁸ +/- la variation des coûts fixés en cette période)) / (le coût moyen des documents existants

⁶ La majorité des méthodes que nous allons citer sont explicitement définies dans tous les livres de Statistiques. Pour le reste, nous donnerons la référence pour un éventuel intérêt.

⁷ Glen J. KELLY.- " Using an Economic Development Approach to Improve Budget Forecasting Techniques, Collection Allocation Methods and Library Budgeting Decision ", In Acquisition Librarian, vol. 5 (1991), pp. 191-213.

⁸ Les coûts relatifs au personnel, la maintenance, l'équipement, ...etc.

+/- la variation du prix moyen des documents) ... etc. ⁹

b) *Modèles basés sur un calcul de probabilité :*

Le principe est d'obtenir, à partir de données brutes, une distribution théorique (exemple : Poisson, Binomiale négative, Gamma, ...) pouvant représenter la distribution réelle de ces données observées.

Bien sûr, une fois trouvée, cette distribution n'est homologuée qu'après l'avoir vérifiée grâce à un des tests de statistique, nous citerons à titre d'exemple le test χ^2 vu la simplicité et l'efficacité de son utilisation.

c) *Modèles basés sur les processus Markoviens :*

Ces modèles sont plus connus en science de l'information sous le nom de Modèles Markoviens de MORSE. Cette confusion est dû au fait que Morse fut le premier à bien utiliser la théorie de MARKOV, dans le domaine de la bibliothéconomie en particulier.

Le principe est que " la demande " suit un processus poissonien (i.e. : la demande suit la loi de Poisson) et que le temps séparant deux demandes consécutives suit la loi exponentielle.

d) *Modèles basés sur la régression linéaire simple (resp. la régression multiple) :*

Le principe est d'établir des prévisions en identifiant un critère (nommé variable) indépendant important (resp. un ensemble de variables), et en spécifiant sa relation (resp. leur relation) avec le critère dépendant que l'on veut estimer.

⁹ Pour plus de modèles voir : Ferdinand F. LEIMKUEHLER, Michael D. COOPER.- " Analytical Models for Library Planning ". In Journal of the American Society for Information Science, November-December 1971, pp. 390-398.

Autrement dit, l'objectif de la régression est de rechercher une relation fonctionnelle entre deux variables (resp. plusieurs).

e) *Modèles basés sur l'étude des séries chronologiques :*

Le principe est d'établir des prévisions à partir de données réelles auxquelles diverses opérations ont appliquées ; exemple : décomposition de la série chronologique, filtrage, ...etc.

f) *Autres modèles :*

Parmi les méthodes non classiques établies, nous citerons deux d'entre-elles car elles sont parmi les plus cotées de celles traitant de la gestion des bibliothèques ; leurs structurations méritent d'être citées ; elles sont assez particulières :

- la méthode qui combine la décomposition d'une série chronologique avec la régression linéaire simple,

- la méthode qui combine la décomposition d'une série chronologique avec la régression multiple.

Remarque : En établissant une comparaison entre ces deux approches, O. N. AHIOKWO ¹⁰ a conclu que la différence n'était pas importante.

3.2. Les méthodes utilisées pour la gestion dans une bibliothèque

A partir de nos lectures, nous avons pu établir une liste des fonctions pour lesquelles les prévisions pour la gestion ont été faites au moyen de certaines méthodes mathématiques que voici :

3.3. La solution au problème de gestion des bibliothèques

Comme il a été signalé, la gestion de la bibliothèque s'est vue utiliser plusieurs méthodes que nous pouvons catégoriser comme suit :

¹⁰ OKECHUKWU N. AHIOKWO.- " Forecasting Techniques and Library Circulation Operations : Implications for Management ". In Library and Information Science Research, vol. 10, n° 2 (1988), pp. 195-210.

- les méthodes analytiques,
- les méthodes probabilistes,
- la régression,
- et les séries chronologiques.

Dans cette partie de l'étude, nous allons tenter de sélectionner l'une des méthodes comme solution à la gestion de la bibliothèque; bien sûr tout en garantissant une capacité à être utilisée pour la majorité, si ce n'est pour la totalité des fonctions à planifier, ainsi qu'une simplicité d'utilisation.

Nous rappellerons qu'il n'est pas nécessaire de valider cette méthode, compte tenu qu'elle appartient à l'ensemble des méthodes mathématiques sus-citées.

Analysons donc chacune de ces méthodes :

- Les méthodes analytiques utilisent des formules mathématiques qui font généralement appel à des fonctions pouvant rendre le calcul difficile. Nous citerons à titre d'exemple le calcul d'intégrale.
- Les méthodes probabilistes, quant à elles, font appel à de bonnes bases en statistiques, surtout lorsqu'il s'agit de calculer les paramètres des lois. De plus, nous savons qu'il ne suffit pas d'obtenir la loi, mais il faut aussi passer par sa validation.

Aussi, ces deux méthodes sont-elles à éviter, compte tenu de l'effort que devra faire le bibliothécaire lors de la gestion de sa bibliothèque, à partir de ces méthodes.

La régression de son côté, malgré sa présence dans chaque logiciel de statistiques, présente des lacunes quant à son utilisation dans la gestion des bibliothèques. En effet, cette méthode ne prend pas en considération les effets cycliques ou périodiques des données (nous citerons à titre d'exemple, les données concernant le prêt des documents).

Par contre, l'analyse des séries chronologiques est non seulement disponible dans les logiciels de statistiques et donc

est simple à utiliser, mais encore elle prend en compte toutes les fluctuations que sont sensées présenter les données.

Ainsi et à priori, cette méthode est candidate. Elle semble répondre à toutes nos exigences. Pour confirmer notre hypothèse, définissons cette méthode.

Définition

Une série chronologique est une suite de valeurs ordonnées dans le temps. Ces valeurs sont réelles.

But

L'analyse des séries chronologiques ou temporelles permet l'étude des phénomènes passés et présents et peut extrapoler l'événement dans le futur. Ainsi donc, la prévision par cette méthode se fonde sur la connaissance du passé et du présent.

Etapes de l'analyse d'une série chronologique

1) La représentation graphique de la série

Cette représentation permet d'entrevoir les quatre composantes fondamentales du mouvement du phénomène étudié, à savoir :

- la composante de tendance (T) : il s'agit de la caractéristique déterminant l'allure générale du phénomène à long terme ¹¹;
- la composante cyclique (C) : il s'agit de la caractéristique déterminant les comportements périodiques dus à des déséquilibres conjoncturels à moyen ou long terme;
- la composante périodique (P) ou saisonnière (S) : il s'agit de la caractéristique déterminant les comportements périodiques, mais cette fois la périodicité dépend de nos activités ou de notre environnement, exemples : les saisons, les fêtes, les congés, ... etc;
- la composante aléatoire (E) ou accidentelle (U) : il s'agit de la caractéristique

¹¹ Dans les prévisions :

court terme ⇔ moins de 3 mois

moyen terme ⇔ entre 3 mois et 2 ans

long terme ⇔ plus de 2 ans.

que résumant les fluctuations accidentelles et donc imprévisibles, exemples : les inondations, les grèves, ...etc.

2) La décomposition de la série

Il s'agit de décomposer la série étudiée en ses principales caractéristiques. Autrement dit, il s'agit de trouver le modèle qui prend en considération l'ensemble des caractéristiques de la série ($X = f(T, C, P, E)$).

Il existe quatre formulations de f :

- le modèle additif : $X = T + C + P + E$
- le modèle multiplicatif : $X = T \cdot C \cdot P \cdot E$
- le modèle mixte : $X = (T + C) \cdot P \cdot E$
- autres modèles : plus ou moins complexes.

Traditionnellement, on retient deux types de modèles : le modèle additif car la tendance y est indépendante des autres composantes et le modèle multiplicatif car toutes les composantes restantes sont proportionnelles à celles de la tendance.

3) L'étude des modèles

Selon le modèle choisi, on fait appel à diverses méthodes d'analyse et de prévision dont les plus utilisées sont :

les méthodes de lissage :

- Parmi les méthodes de lissage, nous citerons : la méthode des moyennes mobiles et la méthode du lissage exponentiel de BROWN.

les méthodes d'ajustement :

- Là, nous citerons la régression.

Après cette brève définition d'une analyse des séries chronologiques, nous allons maintenant nous baser sur les modules communs existants qu'offrent les logiciels de statistiques, pour pouvoir expliciter la procédure d'analyse que devra

suivre le gestionnaire d'une bibliothèque, lors de son étude de prévision.

Du menu principal, nous pouvons déjà localiser : Data, Graphs et Analyses.

Le sous menu Data permet la définition des variables des données de notre étude et donc de nos séries chronologiques.

Le sous menu Graphs permet la représentation graphique de la série étudiée.

Et le sous menu Analyses permet l'analyse des données selon une des méthodes statistiques, en particulier l'analyse des séries chronologiques, en exposant les différentes approches pour le lissage et l'ajustement.

4. COMMENTAIRE FINAL














Il est évident qu'une gestion d'une institution en général et d'une bibliothèque en particulier, basée sur des méthodes et modèles mathématiques rigoureux, demeure la solution la plus efficace.

Dans cette étude, nous avons pu mettre en évidence l'importance de l'analyse à partir d'études de séries chronologiques, dans la gestion des différentes fonctions d'une bibliothèque.

Nous signalerons au passage qu'il existe divers logiciels performants pour une étude statistique tel que définie précédemment, citons : SPSS, SAS, Systat, Minitab et ForeCalc.

Ainsi, la gestion de la bibliothèque repose sur une bonne maîtrise de l'outil informatique et particulièrement celle d'un logiciel de statistiques et non pas sur la maîtrise de la théorie des Statistiques. Il suffira au gestionnaire et/ou bibliothécaire de se former à ce type de logiciel, grâce à une formation spécialisée ou de s'auto-former par le biais du manuel d'utilisation.

BIBLIOGRAPHIE

-  O. D. ANDERSON - *Time series Analysis and Forecasting : The Box-Jenkins approach*, Butterworth & Co, 1976, p. 182.
-  Chen CHING-CHICH - *Applications of operations research Models to Libraries*, The Massachusetts Institute of Technology. The MIT Press, 1976, 212 p.
-  Dennis SMITH - *Forecasting Price Increase Needs for Library materials : the University of California Experience*, In : *Library resource and Technical Services*, 1984, vol. 28, n° 2, pp. 136-148.
-  Ferdinand F. LEIMKUEHLER, Michael D. COOPER - *Analytical Models for Library Planning*, In : *Journal of the American Society for Information Science*, 1971, November-December, pp. 390-398.
-  Georges BRESSON et Alain PIROTTE - *Econometrie des séries temporelles : Théorie et applications*, Presses universitaires de France, 1995, p. 343.
-  Gérard MERCURE - *Les logiciels de prévision et leur utilisation en bibliothèque*, In : *Documentaliste et Bibliothèques*, 1994, juillet-septembre, pp. 155-161.
-  Glen J. KELLY - *Using an Economic Development Approach to Improve Budget Forecasting Techniques*, Collection Allocation Methods and Library Budgeting Decision, In : *Acquisition Librarian*, 1991, vol. 5, pp. 191-213.
-  Linda MAIN - *Computer Simulation and Library Management*, In : *Journal of Information Science*, 1987, vol. 13, n° 5, pp. 285-296.
-  Okechukwu N. AHIOKWO - *Forecasting Techniques and Library Circulation Operations : Implications for Management*, In : *Library and Information Science Research*, 1988, vol. 10, n° 2, pp. 195-210.
-  Roland DUCASSE - *Méthodes du traitement des données bibliométriques pour la gestion des systèmes d'information : Application à l'analyse prévisionnelle de la demande d'ouvrage en bibliothèque*, Thèse de doctorat, Bordeaux, 1978, p. 221.
-  Terrence A. BROOKS - *Using time-series regression to Predict Academic Library Circulation*, In : *College and Research Libraries*, 1984, vol 45, n° 6, pp. 501-505.
-  Terrence A. BROOK and John W. FORYS - *Smoothing Forecasting Methods for Academic library circulation and Evaluation and recommendation*, In : *Library and Information science research*, 1986, vol. 8, n° 1, pp. 29-39.
-  Yves-François LE COADIC - *La science de l'information, Que sais-je ?*, Presses universitaires de France, 1994, p. 127.