

BASVILLE STATISTICIEN SOLITAIRE DU LANGUEDOC

Antoine de Falguerolles

Laboratoire de statistique et probabilités

Université Paul Sabatier

118 route de Narbonne

31062 Toulouse cedex 4

Résumé : Nicolas de Lamoignon de Basville fut intendant du roi Louis XIV en Languedoc. À ce titre, il vécut longtemps à Montpellier. Basville est l'auteur de « mémoires sur la province de Languedoc » rédigés dans les années 1696-1697. Cet ouvrage contient des tableaux statistiques assez élaborés. Deux d'entre eux sont des tables de contingence d'assez grande taille, avec un double système de dénombrement. Ces tables seront plus particulièrement considérées dans cet exposé.

Summary: Nicolas de Lamoignon de Basville (26 April 1648 – 17 May 1724) was the intendant for the King Louis the XIVth in the Languedoc province. As such, he lived in Montpellier for several years. Basville is the author of a « Report on the Languedoc province » written in the years 1696-1697. The report contains a few statistical tables which are quite sophisticated. Two of them happen to be contingency tables of quite large dimensions, with dual counting system. These tables will receive special attention in this presentation.

Mots-clés : table de contingence, SVD complexe, variables aléatoires de Poisson corrélées, histoire de la statistique.

Keywords: contingency table, complex SVD, correlated Poisson random variables, history of statistics.

1 Introduction

Nicolas de Lamoignon de Basville fut intendant du roi Louis XIV en Languedoc. À ce titre il vécut à Montpellier. Durant son intendance, Basville fut l'auteur de « mémoires sur la province de Languedoc » rédigés dans les années 1696-1697. Une des caractéristiques de cet ouvrage est la part qui y est faite aux renseignements chiffrés. En particulier, Basville a rassemblé un certain nombre de tableaux de données assez curieux vu leur ancienneté et leur objet.

L'exemple le plus intéressant est fourni par deux tableaux aux intitulés révélateurs : « état concernant les couvents de filles et le nombre des religieuses qu'ils contiennent . . . » et « carte contenant le nombre d'ecclésiastiques, la quantité des maisons religieuses et couvents de plusieurs ordres, et le nombre des religieux qu'ils contiennent . . . ». Ce sont là des exemples assez anciens de tables de contingence publiées en tant que telles. Mais

il est difficile de trouver dans le texte de Basville l'ébauche d'une analyse statistique au sens où on l'entend actuellement.

C'est l'étude de ces deux tables (« femmes » et « hommes ») d'assez grandes dimensions (23×23 ou 23×25 , selon le tableau) qui est proposée dans cet exposé.

2 Basville

Avant de décrire les tableaux de données rassemblés dans les « mémoires », il peut être utile de présenter brièvement le personnage de Basville.

2.1 L'Intendant du Roi

Nicolas de Lamoignon de Basville (26 avril 1648 – 17 mai 1724), intendant du roi en Languedoc, vécut à Montpellier durant 33 ans. Basville fut en Languedoc (1685–1718), comme auparavant en Poitou (1682–1685), le bras séculier du roi Louis XIV dans l'entreprise de catholicisation forcée des protestants. Aidé d'une trentaine de subdélégués, Basville appliquait la volonté royale sur les 23 diocèses des deux généralités que comprenait son intendance, celle de Montpellier et celle de Toulouse. À ce titre il devait contrôler des zones géographiques à forte densité de nouveaux convertis douteux ou de fanatiques protestants. On sait que, sous son autorité, la répression fut féroce.

Cette présentation du grand commis de l'État que fut Basville est sans doute trop lapidaire et donc forcément injuste. Pour plus de détails sur sa carrière, il faut lire l'ouvrage magistral que le préfet Pierre Pujol a publié en 1992 sous le titre évocateur de « Basville roi solitaire du Languedoc ». (Le titre de mon exposé n'est donc qu'un facile plagiat.) Il faut aussi lire le chapitre III de la première partie de l'édition critique des « mémoires » publiée par Françoise Moreil en 1985. On y verra en particulier que Basville reçut une éducation soignée des jésuites du Collège de Clermont (Paris) où il prononça en 1666 sa thèse « *Ex optica et astronomia selecta mathemata* ». D'une certaine façon, Basville était donc un scientifique.

2.2 Les mémoires

Basville rédigea les « mémoires pour servir à l'histoire de Languedoc » dans les années 1696–1697. Ce travail a été écrit sur ordre du Roi à l'intention du duc de Bourgogne, un de ses proches héritiers. Il fait donc partie d'un ensemble de travaux fournis par les intendants pour la formation et l'information du duc de Bourgogne en réponse à un questionnaire assez précis défini par le duc de Beauvillier.

Les « mémoires » de Basville ont été publiés en 1734, soit dix ans après la mort de leur auteur. Mais ils ont beaucoup circulé sous forme de nombreuses copies manuscrites présentant des différences assez notables selon l'humeur des copistes. Sur les variantes de certaines copies, on se reportera à l'édition critique de Françoise Moreil (1985). Par

exemple, le manuscrit de Marseille (1985, p. 290) comporte un « tableau des feux » croisant diocèses et catégories sociales et donnant aussi le nombre de feux par diocèses. Ce tableau est absent dans l'édition de 1734 ainsi que dans la plupart des manuscrits. À l'étude, il se révèle aussi que les données de cette table sont assez douteuses.

Les « mémoires » comportent de nombreuses données chiffrées qui viennent étayer la présentation et l'argumentation de Basville. Basville a délibérément choisi de répondre aux questions posées par le duc de Beauvillier sous une forme quantitative appuyée. Les données sont soit disséminées dans le texte soit rassemblées sous forme de tableaux statistiques de facture moderne : tableau individus \times variables ou table de contingence. Lorsque l'une des dimensions du tableau de données est géographique (les 23 diocèses des deux généralités de son intendance), Basville désigne parfois ce tableau sous le nom de « carte ».

2.3 Les tableaux de données

Dans l'édition de 1734, on trouve en particulier :

1. Un tableau des « archevêchés et évêchés de Languedoc » (1734, p. 61 et 1985, p.117) donnant pour chacun des 23 diocèses le montant des revenus, le nombre des paroisses, et l'ancienneté documentée du diocèse.
2. Un tableau de l'« état concernant les couvents de filles et le nombre des religieuses qu'ils contiennent ... » (1734, encarté après la page 64 et 1985, p.118-119) donnant pour chaque diocèse (23 lignes) et chaque ordre (23 colonnes), le nombre des couvents et le nombre des religieuses.
3. Un tableau ou « carte contenant le nombre d'ecclésiastiques, la quantité des maisons religieuses et couvents de plusieurs ordres, et le nombre des religieux qu'ils contiennent ... » (1734, encarté avant la page 65 et 1985, p.118-119) donnant pour chaque diocèse (23 lignes) et chaque ordre (24 colonnes) le nombre des couvents ou maisons et le nombre des religieux ; une colonne supplémentaire donne le nombre d'ecclésiastiques. Donc une table de dimension 23×25 .
4. Un tableau de l'« état de la religion à l'égard des nouveaux convertis » (1734, p. 78 et 1985, p.131) donnant par diocèse (23 lignes), le nombre de chefs de famille anciens catholiques (colonne 1), de chefs de famille nouveaux convertis (colonne 2), d'autres habitants anciens catholiques (colonne 3), d'autres habitants nouveaux convertis (colonne 4),
5. Un tableau donnant la « récapitulation du produit du commerce ... » (1734, encarté après la page 300 et 1985, p. 253-256) donnant par type d'activité (49 lignes) l'« estimation de ce qui se fait dans la province » (colonne 1) et l'« estimation de ce qui sort de la province » (colonne 2).

Les tables de contingence 2 et 3 ci-dessus, d'après l'édition de 1734, sont disponibles sous différentes formes sur ma page web : <http://www.lsp.ups-tlse.fr/Fp/Falguerolles/>.

3 Interprétations statistiques

Mais quelles analyses pour ces données et, en particulier, pour les deux grandes tables de contingence reproduites dans ce texte ?

3.1 Basville statisticien

Si l'on ne s'arrête pas à quelques erreurs, souvent typographiques, la lecture des « mémoires » montre que Basville possède une excellente maîtrise de l'emploi des chiffres. Par ailleurs, Basville souligne fort à propos l'intérêt pratique de présenter des données sous forme de tableaux : « pour éviter de même un long et ennuyeux détail des Abbayes et Monastères, je les ai réduit [...] en distinguant ceux des Hommes et des Filles » (p. 64) ou encore « pour connaître tout d'un coup le nombre des nouveaux convertis et les lieux où ils sont établis, il faut jeter les yeux sur la carte ci-derrière » (p. 77).

Si l'aspect quantitatif des mémoires reste assez exceptionnel pour l'époque¹, il n'en reste pas moins qu'on n'y trouve pas de trace de traitement statistique ou de questionnement statistique au sens où on l'entend actuellement.

3.2 Pourquoi s'intéresser aux données de Basville

Les jeux de données anciennes peuvent présenter un intérêt pédagogique. Un débutant peut être sensible à l'expression de l'intuition d'un pionnier. Il peut être aussi motivé par l'environnement décrit par de telles données. Un exemple bien connu est celui fourni par l'article de Davenport et Studdert-Kennedy (1972). Ces auteurs réétudient les données tirées de la « balance des peintres » de Roger de Piles (1758), avec l'ajout d'une variable statistique qualitative. Ils se servent alors de ces données pour illustrer le bon usage de méthodes d'analyses multivariées standards (analyses en composante principale, classification automatique, analyse discriminante, MANOVA ...). De façon humoristique, on peut dire que cet article est sans doute l'archétype d'une analyse de données à la française. D'ailleurs, Venables et Ripley (2003) ne s'y sont pas trompés et ont repris ces données² dans la bibliothèque MASS. Faire, par exemple, dans le logiciel R :

```
> library(MASS); > data(painters).
```

En voyant certaines des données de Basville, et notamment les deux grandes tables de contingence, on ne peut qu'être tenté par une utilisation semblable à celles des données de Piles. Outre la soif d'une inscription éventuelle dans la n^e édition du Panthéon qu'est l'ouvrage de Venables et Ripley, l'intérêt peut être aussi historique. D'une part, l'histoire de l'étude des tableaux de contingence fait état de tables bien postérieures à celles de Basville

¹À titre de comparaison, on pourra se reporter aux « mémoires sur le Languedoc » de Charles Bernard de Ballainvilliers (1788), dernier intendant du Roi à Montpellier. Ces mémoires comportent aussi quelques tableaux de données. Mais il est clair que Ballainvilliers n'a pas la maîtrise des chiffres de Basville.

²Venables et Ripley ont supprimé les unités statistiques présentant des données manquantes.

(Stigler, 2002). Mais les tables sont alors associées à un questionnement de méthodologie statistique. D'autre part, l'historien spécialiste du XVII^e siècle pourrait trouver quelques glanes inédites dans tous ces nombres.

La tâche est cependant difficile à divers titres : d'abord, de par la structure particulière des données ; ensuite, de par la difficulté à modéliser le processus qui les génère.

3.3 Une réanalyse

Basville, toujours méticuleux, sait qu'il faut différents critères de mesure pour quantifier un phénomène. Ainsi, chacune des deux tables de contingence apparie dans chaque cellule le nombre d'établissements et le nombre de religieux. De façon analogue, le tableau de la page 78 des « mémoires » apparie chez les anciens convertis puis les nouveaux convertis le nombre de chefs de famille gentilhommes et le nombre d'autres personnes. Ce sont, là encore, deux mesures du potentiel de résistance à la catholicisation (ou du potentiel de menées subversives conduites par les fanatiques protestants) : rôle des notables, rôle des masses.

Comment modéliser cette redondance d'information qui se prête mal aux traitements statistiques conventionnels. Est-ce le couvent qui fait le religieux ? Ou est-ce le religieux qui fait le couvent ? A l'évidence, ni l'un ni l'autre. Mais comment modéliser cette dynamique ?

Plusieurs techniques peuvent être utilisées pour de telles données appariées. J'en explorerai brièvement deux. Puisqu'il s'agit de dénombrements, une approche de type régression poissonnienne à réponses corrélées peut être envisagée. Mais l'introduction d'un simple coefficient de corrélation entre les nombres de couvents (n) et les nombres de religieux (N) semble bien simpliste pour traduire le mécanisme social sous-jacent. Par ailleurs, la présence de nombreux zéros, résultant encore de ce mécanisme, peut fortement perturber une telle approche (sous-dispersion). Une autre technique, plus descriptive, consiste à faire une analyse factorielle complexe des couples (n, N) codés sous forme de nombres complexes : $n + iN$. On sait que cette approche permet de mettre en évidence synergie et spécificité des deux évaluations (Dossou-Gbété et Grorud, 2002).

Je dois avouer, à ce stade de ma recherche, que les deux approches ci-dessus ne donnent rien de bien exploitable. Les données de Basville sont-elle de trop mauvaise qualité pour dire autre chose que ce qu'il leur fait dire ? Parleront-elles un jour ? Je ferai le point des traitements effectués lors de ma présentation.

Bibliographie

- [1] Ballainvilliers, C. B. (1788) *Mémoires sur le Languedoc*, Entente Bibliophile, Montpellier (réédition de 1989).
- [2] Basville, N. de Lamoignon de (1734) *Mémoires pour servir à l'histoire de Languedoc*, Pierre Boyer, Amsterdam.
- [3] Davenport, M. et Studdert-Kennedy, G. (1972) The statistical analysis of aesthetic judgment: an exploration. *Applied Statistics*, 21(3), 324-333.

- [4] Dossou-Gbété, S. et Grorud, A. (2002) Biplots for matched two-way tables. *Annales de la Faculté des sciences de Toulouse*, Série 6, XI(4), 469–483.
- [5] Moreil, F. (1985) *L'intendance de Languedoc à la fin du XVII^e siècle, édition critique du « mémoire pour l'instruction du duc de Bourgogne »*, CTHS, Paris.
- [6] Piles, R. de (1758) *Cours de peinture par principes*, Jacques Estienne, Paris.
- [7] Poujol, R. (1992) *Basville roi solitaire du Languedoc, intendant à Montpellier de 1685 à 1718*, Les presses du Languedoc, Montpellier.
- [8] Stigler, S. (2002) The missing early history of contingency tables. *Annales de la Faculté des sciences de Toulouse*, Série 6, XI(4), 563–573.
- [9] Venables, W.N. et Ripley, B.D. (2002) *Modern Applied Statistics with S-PLUS. (Fourth Edition)*, Springer.