

FEINSTEIN, Charles H. & THOMAS, Mark, *Making History Count. A Primer in Quantitative Methods for Historians*

Cambridge, Cambridge University Press, 2002, 547 p.

Jean Heffer



Édition électronique

URL : <http://journals.openedition.org/histoiresmesure/1837>
ISSN : 1957-7745

Éditeur

Éditions de l'EHESS

Édition imprimée

Date de publication : 2 décembre 2003
Pagination : 403-408
ISBN : 2-222-96740-6
ISSN : 0982-1783

Référence électronique

Jean Heffer, « FEINSTEIN, Charles H. & THOMAS, Mark, *Making History Count. A Primer in Quantitative Methods for Historians* », *Histoire & mesure* [En ligne], XVIII - 3/4 | 2003, mis en ligne le 05 avril 2007, consulté le 02 mai 2019. URL : <http://journals.openedition.org/histoiresmesure/1837>

Ce document a été généré automatiquement le 2 mai 2019.

© Éditions de l'EHESS

FEINSTEIN, Charles H. & THOMAS, Mark, *Making History Count. A Primer in Quantitative Methods for Historians*

Cambridge, Cambridge University Press, 2002, 547 p.

Jean Heffer

- 1 Pour avoir participé en 1981, en compagnie de Jean-Louis Robert et de Pierre Saly, à la rédaction d'un petit manuel de statistique pour historiens, je n'en suis que plus à l'aise d'affirmer que Charles Feinstein et Mark Thomas viennent de publier le meilleur outil qu'il m'ait été donné de lire. Il est vrai que l'époque a beaucoup changé. En 1981, il n'y avait pas encore de PC et les étudiants investissaient seulement dans des calculatrices limitées aux quatre opérations ; désormais, les ordinateurs personnels abondent et nous délivrent de toute une série de tâches fastidieuses ; il n'est plus nécessaire, par exemple, d'apprendre à utiliser une table de logarithmes. Le manuel de 1981 avait été rédigé pour servir de complément à des travaux dirigés, alors que le livre de Feinstein et Thomas se suffit à lui-même, avec des exercices en fin de chaque chapitre. Enfin, en vingt ans, les techniques statistiques se sont affinées et les historiens ne peuvent plus ignorer ces progrès.
- 2 La première qualité de ce manuel écrit par un professeur émérite d'Oxford, réputé pour ses recherches sur la comptabilité nationale anglaise depuis le XIX^e siècle (Feinstein), et par un brillant professeur associé de l'Université de Virginie (Thomas) est pédagogique. D'une clarté limpide dans la précision des définitions, le texte ne comporte pas de démonstrations mathématiques, à l'exception de brèves notes. Ceux qui désirent approfondir les fondements théoriques des formules sont invités à se reporter à des ouvrages plus spécialisés cités dans la bibliographie, comme ceux de Blalock¹, ou des économètres Kennedy, Maddala, Wonnacott, etc. L'important, c'est de faire comprendre intuitivement ce dont il s'agit et pour cela, nul besoin de barder le texte d'équations qui rebutteraient l'immense majorité des historiens. Seuls, des historiens rompus aux contacts avec des étudiants de cette discipline pouvaient réussir ce tour de force pédagogique, répondant aux questions fréquemment posées dans les classes ou levant les obstacles qui

seraient moins sensibles à un pur statisticien. Feinstein et Thomas partent du principe que l'historien actuel dispose d'un ordinateur et de logiciels du type SPSS, STATA ou SAS, mais ils avertissent à juste titre qu'il ne convient pas de se reposer entièrement sur la machine, qu'il vaut mieux saisir les idées sous-jacentes au calcul de tel ou tel coefficient.

- 3 La seconde vertu de l'ouvrage est son caractère presque complet. Quatre parties, d'une centaine de pages chacune, traitent de l'essentiel de ce qu'il faut savoir pour faire de la bonne histoire quantitative. La cinquième et dernière partie qui rassemble quatre études de cas tirés de la littérature historique anglaise et américaine montre comment des historiens ont utilisé telle ou telle technique statistique en fonction du problème qu'ils avaient à analyser.
- 4 La première partie étudie l'analyse statistique élémentaire. Après avoir défini le vocabulaire de base (population, échantillon, types de variables, logarithmes, indices, trend, fluctuations, etc.), les auteurs présentent la statistique descriptive : les représentations graphiques (en se concentrant sur celles qu'on utilise habituellement), les mesures de tendance centrale, de dispersion, la forme de la distribution, notamment la courbe normale. Les formules apparaissent dans des encarts grisés ; des exemples indiquent clairement la manière dont il faut procéder si l'on ne se contente pas d'appeler la commande adéquate du logiciel. Le chapitre suivant est consacré à la corrélation. Il recommande de partir d'une représentation graphique – ce qui évite les erreurs imputables à une confiance aveugle dans les résultats distillés par la machine. Il est nécessaire de détecter les valeurs extrêmes qui influent sur le calcul. En deux pages, l'essentiel est dit sur l'interprétation du coefficient de corrélation : il ne faut pas confondre corrélation et cause ; un coefficient élevé indique qu'il y a une relation entre deux variables, mais il ne dit rien sur la nature de cette relation ; les valeurs calculées n'ont pas de signification absolue ; un même coefficient sera considéré comme faible ou élevé, selon ce qu'on attend *a priori*. « On doit s'efforcer, lit-on, d'établir des critères jugeant de la force de l'association qui aillent au-delà de la simple transcription des coefficients de corrélation eux-mêmes » (p. 82). Gare aux risques de collinéarité ! De la corrélation, on passe ensuite à la régression linéaire simple, au calcul de la droite de régression, puis à la mesure de la qualité de l'ajustement par le coefficient de détermination. Un développement sur la relation entre le coefficient de régression et le coefficient de corrélation (qui est la racine carrée du coefficient de détermination) me paraît particulièrement bien venu, tout comme le recours à un diagramme inspiré du logo d'une marque de bière (Ballantine) pour faire comprendre la nature du problème même aux plus rétifs des historiens.
- 5 La deuxième partie aborde les échantillons et la statistique inductive, sans entrer dans les fondements mathématiques. Le lecteur est invité à accepter les quatre propositions fondamentales du théorème de limite centrale ; des encarts donnent les indications essentielles à ceux qui voudraient aller un peu plus loin. On trouvera peu de tables (celle du « t » de Student est reproduite pp. 135-136), car les ordinateurs donnent les résultats. Au terme du chapitre, le lecteur est très bien informé sur la distribution de Z et de t, sur les intervalles de confiance pour une moyenne d'échantillon, des coefficients de corrélation ou de régression. Il peut alors aborder le test des hypothèses. Les cinq étapes à suivre sont clairement précisées : d'abord spécifier l'hypothèse sous une forme appropriée pour le test ; choisir ensuite un niveau de probabilité qui décidera de l'acceptation ou du rejet de l'hypothèse, puis une statistique adéquate et sa distribution ; il reste à calculer cette statistique et à la comparer à la valeur théorique pour aboutir à

une décision ; enfin, on interprète les résultats. Ce sont des questions qui exigent une certaine subtilité d'esprit. Feinstein et Thomas réussissent à faire comprendre les deux types d'erreur (type I : rejet d'une hypothèse vraie ; type II : échec du rejet d'une hypothèse fausse), ainsi que la distinction entre tests unilatéral et bilatéral. Avec justesse, ils remarquent que la signification statistique n'implique pas que le résultat soit nécessairement significatif (p. 160). Il ne faut pas confondre signification statistique et signification historique (à savoir ce qui est important du point de vue de l'historien) : un coefficient de régression, par exemple, peut être statistiquement significatif, mais être si faible que la variable exogène n'exerce pratiquement aucune influence sur la variable expliquée. La démarche à suivre est si bien balisée et illustrée par Feinstein et Thomas qu'aucun lecteur ne devrait être perdu en cours de route. La voie est alors ouverte pour s'attaquer au chapitre 7 sur les tests non-paramétriques appliqués à des données qualitatives, comme le sexe (homme/femme) ou la catégorie socio-professionnelle, ou ordinales. On ne peut pas appliquer dans ces cas-là les tests d'hypothèses dont il a été question dans le chapitre précédent. Les auteurs traitent d'abord des tests appliqués à deux échantillons indépendants (Wald-Wolfowitz, Mann-Whitney, Kolmogorov-Smirnov), avec exemples à l'appui et critères pour choisir tel test plutôt que tel autre. Vient ensuite l'inévitable test de χ^2 . Feinstein et Thomas ne négligent pas pour autant les tests portant sur un échantillon unique pour savoir si les observations sont ou non aléatoires. Le parcours s'achève par quatre mesures non-paramétriques de la force de l'association : le ϕ , le V de Cramer, le coefficient de contingence et le t de Goodman et Kruskal, qui sont encore d'un usage plutôt rare.

- 6 La troisième partie est consacrée à la régression linéaire multiple, largement employée par les nouveaux historiens économiques, mais aussi en histoire politique, sociale, voire culturelle. Un dépliant en couleurs représentant des Ballantine (voir plus haut) fait comprendre immédiatement la nature des problèmes lorsqu'on introduit deux variables explicatives au lieu d'une, selon qu'elles sont ou non indépendantes. Pour mesurer l'impact relatif de chaque variable, il est nécessaire de calculer un coefficient *beta* standardisé. Le chapitre 9 peut alors exposer le modèle de régression linéaire classique. La démarche à suivre est clairement indiquée page 259. Il convient de commencer par la spécification du modèle – un modèle explicite, et non implicite comme c'est trop souvent le cas dans les travaux d'historiens – dont les relations peuvent être mesurées et testées. « Est-ce une meilleure approche ? Cela dépend beaucoup de la façon dont les relations sont spécifiées par le chercheur, et de l'exactitude avec laquelle sont mesurées les principales variables. Même si les résultats sont statistiquement pauvres, l'exercice aidera probablement à clarifier les problèmes cruciaux » (p. 259). Les auteurs soulignent à juste titre qu'en matière de modèle, ce qui compte, c'est le jugement du chercheur, plus que la technique statistique proprement dite : il n'y a pas une réponse correcte unique. Les historiens qui se gaussent des querelles entre modélisateurs feraient bien de s'imprégner du passage suivant : « Beaucoup de sujets restent très controversés. Quand des auteurs sont en complet désaccord, les méthodes quantitatives résolvent rarement les disputes. Voilà qui peut surprendre, vu l'apparente précision des mesures et des techniques quantitatives. Mais aucun modèle de relation entre variables ne peut jamais être spécifié précisément et totalement. Il est toujours possible de mettre en question la spécification du modèle et de suggérer d'autres formulations qui pourraient donner davantage de satisfaction » (p. 260). Le chapitre répond parfaitement à son objet : expliquer d'où viennent les différences entre modèles, indiquer les raisons pour lesquelles les critères

d'évaluation aboutissent à des verdicts conflictuels, et montrer pourquoi cependant l'approche est bénéfique, même si les résultats ne correspondent pas à des conclusions définitives. Après avoir rappelé les quatre hypothèses de base du modèle de régression classique, Feinstein et Thomas fournissent tous les outils nécessaires : le test de F pour la signification statistique de la régression dans son ensemble et l'erreur standard de l'estimation. Le chapitre 10 est très important pour les historiens : il porte sur les variables décalées, très utilisées dans le traitement des séries temporelles, ainsi que sur les variables indicatrices (*dummy*). Ces dernières, qui prennent la valeur 0 ou 1, selon que la modalité est ou non présente, sont désormais d'un emploi courant chaque fois qu'on a affaire à des catégories, par exemple, le sexe ou l'océan où l'on pêche la baleine. Les auteurs insistent sur la nécessité de choisir une catégorie de référence qui sera omise de la régression pour éviter le piège qu'on appelle en anglais *dummy variable trap* ; les autres catégories seront estimées par rapport à celle-ci. Les pages 280-289 qui traitent de ces problèmes comptent parmi les plus neuves du manuel, comparé à ce qu'on trouve d'ordinaire dans les ouvrages du même genre destinés aux historiens. Je ne saurais trop en recommander la lecture. La troisième partie s'achève par un chapitre consacré aux limites de la méthode des moindres carrés ordinaires (OLS), quand les hypothèses de base de la régression linéaire classique sont biaisées. Des problèmes sont liés à la spécification du modèle pour cause de non-linéarité, de variables omises ou redondantes, ou de valeurs instables des paramètres. D'autres sont imputables à la spécification du terme d'erreur, qu'il s'agisse d'hétéroscédasticité ou d'autocorrélation des résidus. D'autres, enfin, relèvent de la spécification de la variable, notamment de la simultanéité et de la multicollinéarité. Les auteurs proposent des remèdes ou des stratégies pour débusquer cette dernière, mais sans trop d'illusion. Ils concluent : « La multicollinéarité doit être acceptée comme un risque professionnel par ceux qui calculent des régressions. En un sens, son existence... indique la limite des méthodes quantitatives et souligne que le calcul de régressions n'est pas un substitut au raisonnement historique et à la construction de modèles adéquats et logiques » (p. 323).

- 7 La quatrième partie approfondit l'analyse de régression en étudiant les modèles non-linéaires et les modèles *logit* ou *probit*. Certains modèles non-linéaires peuvent s'estimer par les moindres carrés ordinaires ; dans le cas d'une parabole, il suffit de remplacer la variable χ^2 par une nouvelle variable Z linéaire ; plus généralement, on transforme les variables expliquée et explicatives en logarithmes ; pour les séries temporelles, on utilise plus volontiers un modèle semi-logarithmique. C'est dans ce chapitre 12 que Feinstein et Thomas reviennent sur l'ajustement des séries temporelles, déjà abordé dans le premier chapitre, en prodiguant de judicieux conseils appuyés sur des exemples tirés d'une littérature historique fort variée. Ce sont des choses bien connues des praticiens. À l'inverse, le chapitre suivant est une réelle nouveauté. À ma connaissance, aucun manuel de statistique pour historiens n'avait encore abordé les modèles *logit*, *probit* et *tobit*. Comme les deux derniers sont proches du premier, il est normal de concentrer l'analyse sur celui-ci. De plus en plus, les recherches un peu poussées font appel à cette spécification. En effet, souvent on traite de variables dépendantes discrètes (non continues). C'est le cas avec les variables *dummy*, ou avec celles qui ne peuvent varier qu'entre 0 et 1 (en pourcentage, entre 0 et 100 %), ou bien encore avec les variables censurées (qui n'admettent pas de valeurs négatives, par exemple). Le modèle linéaire des moindres carrés ordinaires ne convient plus pour de telles estimations ; une courbe logistique en S est un meilleur ajustement qu'une droite. Le modèle *logit* est un modèle de régression qui suppose une distribution logistique, le modèle *probit* est fondé sur une

distribution normale. Feinstein et Thomas guident le lecteur pas à pas et lui font découvrir toutes les subtilités de la démarche depuis l'approche du maximum de vraisemblance jusqu'à l'interprétation des résultats fournis par l'ordinateur. La distinction entre l'effet marginal et l'effet d'impact (le premier concernant l'influence d'une variation infinitésimale de la variable explicative, le second la variation relative à une variation finie de cette même variable) est d'une clarté tout-à-fait remarquable, tout comme la différence entre l'élasticité de point et l'élasticité unitaire. Ce chapitre 13, à lui seul, représente un pur joyau et rend le manuel indispensable.

- 8 Devant un si beau travail, on aurait mauvaise grâce à chipoter. Dans un manuel de statistique, on ne sait pas généralement où placer le chapitre sur les indices, car il ne fait partie ni de la statistique descriptive, ni de la statistique inductive. Feinstein et Thomas ont trouvé une solution élégante : ils le mettent en annexe (pp. 507-525). Leur livre garde ainsi l'avantage d'être centré logiquement sur la statistique descriptive et surtout inductive. Mais les méthodes quantitatives (ou la cuisine quantitative) couvrent un périmètre plus vaste. On ne trouvera pas dans le manuel, par exemple, les formules des indices de concentration, comme celui de Gini ; il est vrai que les ordinateurs les donnent aujourd'hui, sans qu'on ait besoin de se fatiguer à les calculer. Le lecteur français s'étonnera de l'absence des autres formes d'analyse multivariée : analyse en composantes principales, analyse factorielle des correspondances (AFC), analyse de corrélation canonique, analyse discriminante, etc. Dans le manuel de 1981, Jean-Louis Robert avait consacré un chapitre de 27 pages (pp. 223-249) à l'analyse des données, notamment à l'AFC, qui résumait très clairement la démarche. Il est vrai que, si l'on voulait présenter ces types d'analyse avec autant de minutie pédagogique que Feinstein et Thomas l'ont fait pour la régression, le volume, qui dépasse 500 pages, devrait en compter au minimum 200 de plus – ce qui est beaucoup pour un cours supposé être suivi en une année universitaire. On peut donc comprendre pourquoi les auteurs ont fait le choix de laisser de côté tout ce pan de l'analyse multivariée. J'ajouterai que les historiens « anglo-saxons » ont de bonnes raisons d'être sceptiques à l'égard de l'AFC qui fut considérée en son temps comme une autre exception culturelle française. Le comité de rédaction d'*Histoire & Mesure* ne manque pas de thuriféraires de l'AFC, à laquelle le numéro XII, 3/4 de 1997 a été consacré. Pour ma part, je partage entièrement les conclusions d'Hervé Le Bras exposées dans ce même numéro, dans son article intitulé : « L'analyse multivariée est-elle une méthode générale de recherche ? ». Selon la réponse qu'on apporte à cette angoissante question, on souffrira d'un manque à la lecture du manuel de Feinstein et Thomas, ou l'on aura l'impression d'accéder à l'ouvrage de référence moderne dont on avait besoin et que toute bibliothèque, universitaire et de chercheur, devrait posséder. Je me range délibérément dans cette seconde catégorie.

NOTES

1. BLALOCK, *Social Statistics*, 1979.