

UM EXEMPLO DE UTILIZAÇÃO DE UMA ANÁLISE IMPLICATIVA PARA O EXAME DE QUESTIONÁRIOS

Régis Gras*
Annie Larher**

Problemática didática

A didática, tanto do lado do professor como do pesquisador e excluindo-se algumas áreas, atualmente não dispõe de respostas categóricas com relação às questões que lhe são formuladas. Ora, para adquirir fundamento científico, ultrapassando a simples opinião, ela deve poder formular e antecipar hipóteses correspondentes a essas questões. Ela deve poder estabelecer um dispositivo confiável e fiel de compilação e tratamento de dados capaz de corroborar ou invalidar as hipóteses e levar a conclusões. Certamente essa estratégia ambiciosa mas rigorosa não pode ser montada desde os primeiros contatos com os fenômenos a serem observados e de onde surgem as questões. Ela se impõe, entretanto, posteriormente, se desejarmos que as decisões didáticas se apoiem em regularidades, em uma estabilidade e uma pertinência de respostas e obtenham assim adequação e validade.

Por exemplo, por meio de uma análise apriorística de uma situação-problema, imaginamos a existência de certos procedimentos de resolução e uma hierarquia de eficácia. A observação leva ao aparecimento de um desvio entre o modelo *a priori* e o conjunto dos procedimentos efetivamente observados. Quais conclusões podemos tirar dessa distorção?

* Do I.R.M.A.R. Universidade de Rennes e do I.R.E.S.T.E., Universidade de Nantes.

** Do I.R.M.A.R., Universidade de Rennes.

Outro exemplo: um questionário sobre uma parte heterogênea de um programa escolar é apresentado a uma vasta população de alunos. Supomos complexidades específicas: a independência de certos itens, de certos campos do saber, competências particulares a certas famílias de alunos, etc. A observação evidencia a existência de certos fatores discriminantes. Quais são eles? Qual a sua hierarquia efetiva? Como se posicionam as famílias de alunos em relação a eles?

Outro exemplo: uma vez observadas estratégias de resolução de problemas, podemos associá-las a concepções consistentes? São elas evolutivas e como?

Último exemplo: em psicologia cognitiva, será possível evidenciar, por meio de um questionário adaptado, um segmento de uma epistemologia genética diferencial?

Surgem dificuldades a todo momento e o pesquisador isolado encontra-se desarmado diante das alternativas que se apresentam e das decisões a serem tomadas. Por exemplo, como processar as informações quantitativas? Como codificar os dados? A partir de que número de alunos assegura-se a credibilidade de um resultado? Qual método estatístico podemos adotar? Como interpretar os resultados? Trata-se de encontrar um equilíbrio adequado, na busca da validação de hipóteses, entre a depreciação dos métodos estatísticos, a recusa de investimento nessa área e a "estatisticomania" que conduz a um excesso de resultados inutilizáveis, acompanhado da ilusão de transparência.

Ruptura epistemológica da estatística clássica: a análise dos dados, suas possibilidades de resposta

Uma dupla conjuntura possibilitará respostas satisfatórias à nossa problemática:

— de um lado, a formalização da álgebra linear, da geometria e das probabilidades possibilitará a elaboração de novos métodos de tratamento de dados;

— de outro lado, o computador permitirá acumular esses dados, efetuar cálculos sobre estruturas complexas sem alterar a extensão dos quadros a serem tratados e fornecer representações variadas da informação obtida.

Com efeito, a *análise dos dados — metodologia* de tratamento dos dados visando à modelização dos fenômenos — fornece atualmente múltiplos *métodos*, chamados *análises de dados*, que permitirão obter, contrariamente à sua designação, *sínteses* desses dados, em uma visão holográfica, dos *fatores* discriminantes, das tipologias, hierarquias, etc.

Assim, a ruptura epistemológica diz respeito tanto aos *objetivos* esperados e atingidos, aos *meios técnicos* para obtê-los (informática), aos *dados* tratados (número, natureza, variedade, etc), aos *modos de restituição* da informação, aos *procedimentos* (ir dos dados em direção aos modelos e não o inverso), aos *métodos* matemáticos empregados e aos *conceitos* neles implícitos, etc. Neste sentido, a análise dos dados distingue-se tanto da estatística de inferência e decisão como da estatística descritiva.

Porém, as novas perspectivas oferecidas, aparentemente com um entusiasmo sem reservas, criam ou mantêm a ficção de que dados colhidos e tratados sem escolha adequada do método e sem hipóteses preliminares fornecerão informações claras e resultados organizados. Muitos pesquisadores, que por sinal abandonaram essa metodologia em seguida por este motivo, terminaram com pilhas de papel não utilizável. Desperdício econômico e intelectual! Parece-me indispensável, cerca de vinte anos após os meus primeiros contatos com a análise de dados, proceder da seguinte forma:

— formular hipóteses sem alimentar a ilusão de que elas possam ser refutadas ou definitivamente aceitas, mas somente questionadas ou corroboradas;

— escolher um método de análise adaptado; por exemplo, se buscamos evidenciar:

- os principais fatores discriminantes em uma população por meio de variáveis: uma análise fatorial;
- uma divisão entre as variáveis: as populações dinâmicas;
- uma tipologia ou uma classificação: uma classificação hierárquica das similaridades;
- uma implicação entre variáveis ou classes de variáveis: uma árvore implicativa ou uma hierarquia implicativa, etc.

— conhecer de forma sucinta os conceitos matemáticos que embasam as sínteses (distância do x^2 , por exemplo); esse conhecimento controlará e facilitará a interpretação;

— interpretar os resultados numéricos e gráficos de forma sintética adotando um certo distanciamento, e saber ampliar ou restringir os dados para a realização de uma segunda passagem que pareça necessária para confirmar ou criticar as primeiras interpretações.

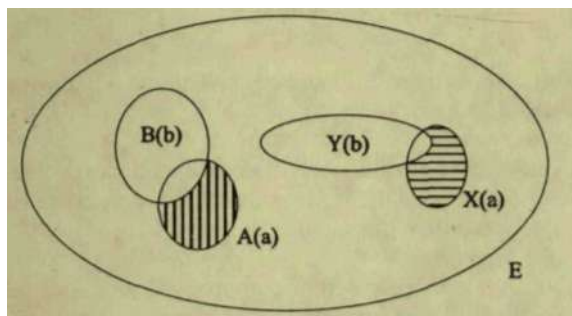
Eventualmente, neste caso, utilizar um método de inferência.

Será necessário, para o pesquisador, ultrapassar as evidências de certos resultados, servindo-se dessa convenção para dar crédito às interpretações mais ocultas, mais surpreendentes, que justifiquem por si próprias o emprego de um método sofisticado.

A análise implicativa

Implicação entre variáveis

Contrariamente aos métodos mencionados acima, onde distância e índice de similaridade são simétricos, o método implicativo, que criamos e desenvolvemos, é não-simétrico. A problemática que o introduz é a seguinte, no caso onde as variáveis consideradas sejam binárias (um indivíduo satisfaz ou não uma variável):



Se A e B são as subpopulações dos indivíduos que satisfizeram as variáveis a e b respectivamente, até que ponto podemos dizer: "se a então b". uma vez que a implicação não deva ser conotada *a priori* de causalidade.

Se $A \subseteq B$, a proposição é verificada; mas, geralmente, os casos mais freqüentes apresentam uma intersecção $A \cap B$ não vazia.

O índice de implicação mede, de forma comparável à similaridade de I.C. Lerman, o grau de "estranheza" diante da pequenez de $A \cap B$, tendo em vista a independência *a priori* e os efetivos observados. Desta forma, diremos, por exemplo, que X e Y sendo duas partes aleatórias de E de mesmos cardinais respectivos que A e B e descrevendo de maneira independente a totalidade das partes de E:

"a \Rightarrow b" é admissível no nível de confiança ou com a intensidade implicativa de 0,95 se e somente se:

$$\text{Prob} [\text{card}(X \cap Y) < \text{card}(A \cap B)] < 0,05.$$

Esta noção foi ampliada, desde a tese de A. Larher, em variáveis modais e numéricas, unificadas em variáveis freqüenciais.

Uma árvore implicativa capta a ordem parcial induzida por essa intensidade de implicação. Ag Almouloud (1992), Ratsimba-Rajohn (1992) e Totohasina (1992) evidenciam em sua tese uma ordem parcial entre os procedimentos empregados por estudantes no tratamento de exercícios, procedimentos inclusos na definição de concepções ou de modelos mais ou menos funcionais. Eles enfatizam as contribuições respectivas dos métodos de análise que empregam. Londeix, em sua tese, reconstrói, a partir de um teste, uma hierarquia de estágios segundo Piaget e mostra uma defasagem diferencial devido aos contextos dos exercícios do teste.

Implicação entre classes de variáveis

Insuficientemente sintética, a implicação entre variáveis é conceitualmente extensível a uma implicação entre classes de variáveis. Uma vez que o exame de tal relação entre duas classes não tem sentido verdadeiro exceto no caso de um "bom fechamento" das classes, definimos o conceito de *coesão* de uma classe como oposto ao de "desordem implicativa" (no sentido da entropia na teoria da informação). Desta forma, a implicação entre duas classes bem "coesivas", isto é, já ordenadas em seu interior, traduz a força implicativa de uma sobre a outra.

Um exemplo de análise em Geometria (A. Larher, 1991)

Observações e alguns estudos mais aprofundados de produções de alunos, de 12-14 anos principalmente, sobre os problemas com demonstração geométrica, mostraram a grande quantidade e a grande variedade de

procedimentos errôneos dos alunos, apesar da estrutura da solução já estar esclarecida. Certamente, os erros têm origem profunda na ausência de significação da prova matemática e em uma certa carência de domínio do léxico necessário (*já que, portanto, pois, porque...*), mas também de forma conseqüente ou conjunta:

— na ausência de rigor na articulação dissimétrica dos três elementos-chave da inferência: hipótese - teorema - conclusão;

— na aceitação de indicadores extrínsecos para a escolha de qualquer um desses elementos-chave:

- indicadores formais (estrutura, ritmo ...);
- indicadores semióticos (palavra, letra, símbolo...);
- indicadores semânticos (um sentido próximo, uma utilização anterior...).

Todo professor sabe bem como é difícil, e até mesmo impossível, identificar em cada trabalho escolar o tipo de erro cometido e sobretudo a sua repetição pelo aluno, sua frequência na classe e as condições nas quais o erro se forma e aparece. Além disso, é ainda mais difícil para ele descobrir para cada aluno as situações que permitiriam tornar esses procedimentos — ou até concepções — conscientes e desequilibrá-los. O computador, em compensação, possibilita um trabalho mais individualizado e, sobretudo uma sanção imediata do erro e portanto que o aluno faça uma revisão de seus procedimentos.

Metodologia adotada

Parece portanto importante, para em seguida trabalhar melhor esses procedimentos com cada aluno, identificá-los e reconhecer suas circunstâncias de aparecimento.

Nossa missão didática e informática¹ consistirá, então, a mais ou menos longo prazo:

— em construir situações onde as variáveis são controláveis;

— em identificar e interpretar os erros e as condições de seu aparecimento;

— em construir um modelo de previsão de procedimentos errôneos:

— em elaborar programas computacionais que satisfaçam os objetivos didáticos.

Esquemáticamente, levando em conta esses objetivos, o microcomputador é integrado sob dois aspectos:

— auxílio didático ao aluno em uma situação de problema com demonstração (*software* DEFI: "Demonstração e Análise da Figura Interativa" que nós desenvolvemos);

— auxílio ao professor para melhor compreensão dos erros cometidos pelo aluno e se possível, sua correção (*software* apresentado mais adiante).

Parece necessário limitar as variáveis em interação em uma demonstração, fornecendo ao aluno situações onde o nexo desenvolvido para o objetivo distante dessa demonstração não seja o estímulo essencial e onde o léxico seja reduzido.

Para tanto, estabeleceremos uma lista de fatos matemáticos (geométricos, no caso) que possam ensinar, de acordo com a situação, hipóteses ou conclusões e uma lista de teoremas. Tendo sido proposta uma inferência incompleta (ou mesmo um problema com demonstração), o aluno deverá, de forma pertinente, escolher um ou vários fatos, um ou vários teoremas para que a inferência ou as inferências sucessivas sejam validadas.

¹ No contexto do Grupo de Pesquisas do C.N.R.S.: "Didática e aquisição de conhecimentos científicos".

O trabalho do aluno será executado em microcomputador, com o auxílio de um *software* que possibilita um trabalho personalizado, seguido de uma análise individual de suas respostas (após eventualmente duas tentativas).

Esse *software* não é um "tutorial" propriamente dito, mas, mais exatamente, uma ferramenta de diagnóstico com três funções:

- reforço do aprendizado das regras de dedução em uma etapa;
- balanço das aquisições sobre o ponto precedente;
- revelação e meio de análise dos erros por estudo diagnóstico

Dependendo da escolha feita no início pelo professor, o aluno terá diversas chances ou não e a resposta correta lhe será ou não fornecida.

Apresentação do questionário

Um conjunto de seis questões é portanto proposto a alunos do segundo ano do colégio (12-13 anos) após o ensino de algumas propriedades da simetria em relação a um ponto. A cada questão corresponde uma inferência que o aluno deve completar escolhendo um dos 11 fatos seguintes a título de conclusão:

Fatos

- | | |
|----|--|
| 1 | (EF) e (CD) são simétricos em relação ao ponto I |
| 2 | [MN] é o simétrico de [PR] em relação ao ponto I |
| 3 | (AB) e (CD) são simétricos em relação ao ponto O |
| 4 | (MN) // (PR) |
| 5 | (CD) // (EF) |
| 6 | (AB) // (CD) |
| 7 | (AB) // (EF) |
| 8 | MN = PR |
| 9 | CD = EF |
| 10 | AB = CD |
| 11 | AB = EF |

Teoremas

- | | |
|---|---|
| 1 | A simetria central conserva os comprimentos. |
| 2 | Se (D) // (D') e (D') // (D''), então (D) // (D''). |
| 3 | A simétrica de uma reta (D) em relação a um ponto é uma reta (D') // (D). |
| 4 | Se duas retas são simétricas em relação a um ponto, então elas são paralelas. |
| 5 | Dois segmentos simétricos em relação a um ponto têm o mesmo comprimento. |
| 6 | A simetria central conserva as direções. |

Questão: Dados hipótese e teorema das listas acima, encontrar a conclusão tirada da lista dos fatos (duas tentativas possíveis para cada questão).

Demonstrações

	HIPÓTESES	TEOREMA	? CONCLUSÃO a encontrar
Q1	Hipótese: 1 Teorema: 3 Conclusão: 3	(EF) e (CD) simétricos em relação a I	(EF) // (CD)
Q2	Hipótese: 4 Teorema: 4 Conclusão: 6	(AB)e(CD) simétricos em relação a O	(AB) //(CD)
Q3	Hipótese: 2 Teorema: 5 Conclusão: 8	(MN) e simétrico de (PR) em relação a I	MN = PR
Q4	Hipótese: 3 Teorema: 6 Conclusão: 6	(AB)e(CD) simétricos em relação a O	(AB) //(CD)
QS	Hipótese: 6e 5 Teorema: 2 Conclusão: 7	(AB)//(CD)	(AB)//(EF)
<J6	Hipótese: 2 Teorema: I Conclusão: 8	(CD) // (EF) [MN] e simétrico de (PR) em relação a I	MN = PR

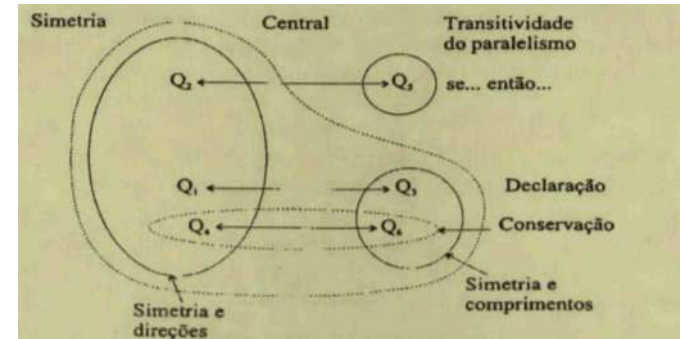
Por meio do questionário, buscamos controlar o efeito das seguintes variáveis didáticas:

- o conceito: 5 das 6 questões tratam da simetria central; uma questão trata da transitividade do paralelismo (questão nº 5);
- a especificação ou competência² dos teoremas (exemplo: teorema 1 *versus* teorema 4);
- grau de generalidade da invariante relacionai (exemplo: teorema 1 *versus* teorema 5);
- a complexidade léxica (exemplo: "conservar") ou conceituai (exemplo: "direção");
- a formulação "se ... então" (teorema 2);
- a simetria da relação entre os objetos denominados (exemplo: fato 1 *versus* fato 2);
- a confusão entre // e = (exemplo: fato 4 *versus* fato 8);
- a expressão de propriedades (exemplo: teorema 1);

Uma análise apriorística da complexidade nos leva a prever as seguintes hierarquias entre os sucessos **Si : $S_2 > S_1 > S_4$; $S_3 > S_6$; $S_2 \geq S_3$; $S_6 > S_4$.**

Podemos esquematizar as proximidades *formal*, *semântica* e *referencial*, *a priori*, dessas seis questões:

Do francês *instanciation* (N.Trad.).



Efetuamos para esse questionário o tratamento estatístico dos dados colhidos conforme dois métodos de análise: a classificação hierárquica (segundo Lerman) e a classificação implicativa (segundo Gras). Veremos mais adiante os resultados que deduzimos do segundo método.

Podemos nos perguntar desde já quais são as bases da estratégia de decisão do aluno neste exercício muito especial que consiste em fazer uma escolha entre um conjunto fechado de soluções. Ela é necessariamente muito próxima da estratégia desenvolvida nas Questões de Múltipla Escolha e, em compensação, muito diferente da que é seguida nas demonstrações em várias etapas, nos problemas apresentados e mesmo no *software* DEFI. Aqui, o aluno deve somente aceitar ou rejeitar um elemento de uma lista. Ele não tem atividade criativa verdadeira. Além disso, o sentido global não é acessível; os únicos pontos de apoio são o sentido da etapa de demonstração e o conjunto léxico das assertivas ou teoremas de que dispõe o aluno. No entanto, observamos, graças à repetição, à acumulação e a concomitância de erros, a estabilidade de certos procedimentos que correspondem a modelos de funcionamento em equilíbrio tanto em um aluno em particular como em alunos em geral. Os erros, que geralmente chamamos de "erros de raciocínio" procedem de causas profundamente arraigadas e não necessariamente de ordem lógica. Devem-se igualmente ao conhecimento errôneo dos objetos tratados (quando não do vocabulário utilizado), e também, com muita frequência, ao poder de atração de certas palavras, certos sinais ou símbolos, certas formas (estruturas de frases, ritmos...) quando da articulação *hipótese*—*teorema*—>*conclusão*.

O aluno se confunde mais, quando erra, diante de um critério "sinal" do que de um critério "sentido". Ele vai buscar nas soluções oferecidas os indícios formais mais plausíveis, os que lhe parecem mais pertinentes.

Resultados: parâmetros dos sucessos

Médias

Reencontramos a hierarquia presumida *a priori* entre os sucessos *S*. nas seis questões: S_2 (96,25%), S_5 (78,75%) e S_4 (72,5%).

Além disso: $S_3 = S_6$ (87,5%).

A taxa de sucesso de Q_5 (85%) é um pouco inferior à taxa de sucesso de Q_5 e Q_6 (Q_5 não faz referência à simetria central; seu teorema tem competência diversa dos outros³. Ela é claramente inferior à de Q , apesar da mesma formulação do teorema em "se... então..."; será por causa da dupla hipótese?

Coeficientes de Correlação entre as Modalidade "de Sucesso" das seis Questões

As mais fortes relações positivas são observadas entre:

S_1 e S_5 (formulação diferente do teorema mas o mesmo conteúdo): $P = 0,38$;

S_1 e S_6 ($p = 0,358$): trata-se dos mesmos alunos que têm dificuldade para começar (Q_1) e para concentrar a atenção (Q_6)?

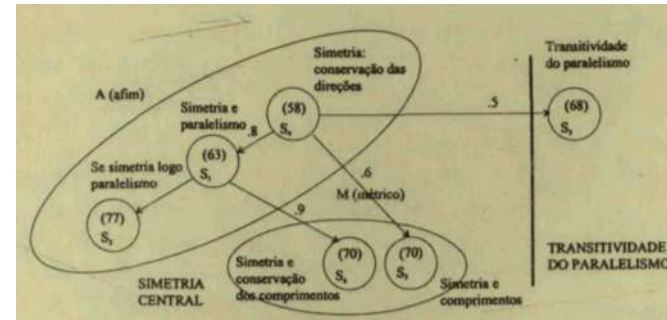
S_3 e S_5 têm um coeficiente de correlação bem próximo de 0, e mesmo negativo, com todos os outros sucessos exceto com S_4

³ Do francês *son ihéortme est instancié* (N.Infd.).

Análise implicativa dos seis sucessos

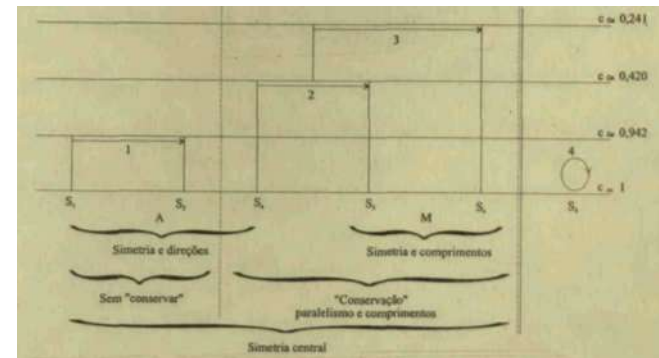
O quadro das implicações permite construir o gráfico orientado transitivo, ponderado e associado à relação de quase-implicação.

Árvore Implicativa de Sucessos



Hierarquia Implicativa

Retomando o método desenvolvido por R. Gras e A. Larher, podemos constituir classes de sucessos que se organizam dessa forma em função das implicações intraclasses.



Encontramos similaridades bastante surpreendentes com as classes formadas *a priori* a partir das proximidades formal, semântica e referencial das *seis* questões:

— separação bem nítida de S_5 , sucesso na única questão relativa à transitividade do paralelismo;

— classe (S_1S_2): as questões Q_1 e Q_2 diferenciam-se apenas pelas expressões de seus teoremas: nenhum deles contém a palavra "conservar" de compreensão ambígua e, de toda forma, difícil para os alunos;

— classe (S_4, S_3, S_6) que reagrupa os sucessos nas duas questões (Q_3) e (Q_6) relativas à propriedade métrica da simetria central e do sucesso na questão (Q_4). de natureza afim mas cujo teorema, como o de (Q_6), é expresso em termos de conservação. Esse último ponto colocará os 3 itens em um mesmo nível de complexidade?

Concluindo, haverá necessidade de indicar o poder de exploração desse método de análise dos fatos didáticos? Constatamos que ele possibilita e possibilitará, graças à sua implementação em informática, decisões em tempo real a partir de estratégias, e mesmo de concepções, extraídas da concomitância repetida de comportamentos. Ela se impõe como ferramenta nova e complementação dos métodos simétricos, oferecendo às questões didáticas as hipóteses de estabilidade que elas buscam.

Referências bibliográficas

AG ALMOULOU, S. *L'ordinateur, outil d'aide à l'apprentissage de la démonstration et de traitement des données didactiques*. [S.1.], 1992. Tese (Doutorado) — I.R.M.A.R., Université de Rennes 1.

GRAS, R. *Contribution à l'étude expérimentale et à l'analyse de certaines acquisitions cognitives et de certains objectifs didactiques en mathématiques*. Tese (Pós-Doutorado) — Université de Rennes 1.

_____. Data analysis: a method for the processing of didactic questions. *Research in Didactic of Mathematics*, Grenoble, v. 12, n. 1, p.59-72, [19-].

GRAS, R., LARHER, A. Uimplication statistique, une nouvelle méthode d'analyse des données. *Mathématique, Informatique et Sciences Humaines*, n.120, 1992.

GRAS, R., TOTOHASINA, A., AG ALMOULOU, S., RATSIMBA-RAJOHN, H., BAILLEUL, M. *La méthode d'analyse implicative en didactique: applications*. Grenoble: La Pensée Sauvage, 1994. R.D.M., 14/1.

LARHER, A. *Implication statistique et applications à l'analyse de démarche de preuve mathématique*. [S.1.], 1991. Tese (Doutorado) — I.R.M.A.R., Université de Rennes 1.

LERMAN, I.C., GRAS, R., ROSTAM, H. Elaboration et evaluation d'un indice d'implication pour des données binaires. *Mathématiques et Sciences Humaines*, n.74/75, 1981.

RATSIMBA-RAJOHN, H. *Contribution à l'étude de la hiérarchie implicative: application à l'analyse de la gestion didactique des Phénomènes d'ostention et de contradiction*. [S.1.], 1992. Tese (Doutorado) — I.R.M.A.R., Université de Rennes 1.

TOTOHASINA, A. *Méthode implicative en analyse de données et application à l'analyse de Conceptions d'étudiants sur la notion de probabilité conditionnelle*. [S.1.], 1992. Tese (Doutorado) — Université de Rennes 1.