

Modelagem Difusa da Mortalidade Infantil pelo Produto Interno Bruto e Cobertura de Sarampo

Milton Urrutia ¹, Ernesto Araujo ^{1,2,3}

¹ Departamento de Informática em Saúde (DIS), Universidade Federal de São Paulo (UNIFESP), Brasil

² Hospital Municipal Dr. José de Carvalho Florence, Brasil

³ Laboratório de Integração e Teste (LIT), Instituto Nacional de Pesquisas Espaciais (INPE), Brasil

Resumo – As relações da mortalidade infantil com o produto interno bruto per capita e a cobertura urbano-rural da imunização de sarampo entre menores de um ano é obtido pela modelagem difusa neste artigo. A fim de gerar uma base de conhecimento os algoritmos de aprendizado Wang & Mendel e Gath-Geva são aplicados a um conjunto de dados fornecidos pela Organização Mundial de Saúde (OMS). A eficiência destes algoritmos para a geração do conjunto de regras e das funções de pertinência é, em seguida, verificada. Resultados preliminares indicam que as relações encontradas coincidem com as achadas na literatura e apresentam evidências para discutir os aspectos controversos do significado da mortalidade infantil. O uso da lógica difusa é um complemento para aprimorar ou descartar relações quando dados são imperfeitos, e como uma ferramenta efetiva para explorar significados de variáveis quando uma base de dados está disponível.

Palavras-chave: Mortalidade Infantil, Lógica difusa, Indicador de Saúde, Wang & Mendel, Gath-Geva .

Abstract – The relationship of infant mortality with gross national income per capita and measles immunization coverage among one-year-olds ratio urban-rural is achieved by fuzzy modeling in this paper. In order to generate the knowledge base, the Wang & Mendel and Gath-Geva learning algorithms are applied to a database supplied by the World Health Organization. The efficiency of these algorithms for generating the set of rules and the membership functions are, then, verified. Results show that the achieved relationships coincide with the literature and present evidence to argue against the controversial aspects of what means infant mortality. The use of fuzzy logic is a complement for consolidate or discard relationships when data are imperfect and it is an effective tool to explore the meanings of variables when database are available.

Key-words: Infant Mortality, Fuzzy Logic, Health Indicator, Wang & Mendel, Gath-Geva.

Introdução

O entendimento das relações da mortalidade e indicadores econômicos data da década de 60. Padrões para o produto interno bruto, urbanização, fertilidade, cuidado médico, iniquidade e educação feminina são variáveis que influenciam na taxa de mortalidade infantil no período de 1950-1975 que envolveram diversos países [1]. Adicionalmente, pode se considerar que a mortalidade infantil tem uma tendência secular que é afetada por fenômenos naturais e guerras e, em seguida, volta a sua tendência [2]. Neste sentido, dois seriam os fatores que a determinam e funcionam como insumos que se acumulam e valorizam no tempo: a infra-estrutura física e o desenvolvimento social dos países. Ainda, há a vertente que doenças epidêmicas devem descrever uma parte da mortalidade que não é possível relacioná-la com alguns dos fatores até aqui descritos [3].

A iniquidade na distribuição do produto interno bruto em países industrializados também tem sido relacionada com a variação no nível médio da saúde da população e que a

comparação entre países não é completamente possível, e é uma matéria de grau [4]. Todavia, este argumento, baseado em trabalhos que fazem uma associação mais forte da iniquidade com a saúde da população, apresentam problemas metodológicos e uma pobre qualidade de dados que debilitam a inferência. Um problema é que a iniquidade pode ser calculada de diversas maneiras pelo coeficiente de GINI e há poucos dados para obtê-lo, sem contar que geralmente se requer ajuste. Idealizado pelo italiano *Corrado Gini*, o coeficiente que leva seu nome, é uma medida de desigualdade do produto interno bruto e tem sido uma ferramenta importante para análise de parâmetros econômicos. Estudos têm propostos modelos de regressão multivariada para mortalidade infantil quando comparado com produto interno bruto e GINI [5,6]. Estes estudos apontam diferenças importantes para coeficiente de correlação e para coeficiente de GINI, respectivamente, de 0,9 e 0,2, aproximadamente. Sustentam ainda que países com baixo nível de produto interno bruto podem ter uma baixa taxa de mortalidade infantil, visto que apresentam um baixo nível de

iniqüidade. A consistência deste argumento é defendida quando se sustenta que as diferenças na distribuição do produto interno bruto estão associados a padrões de saúde na população [6]. Desta maneira, a iniquidade deve estar relacionada com as diferenças das classes sociais para ser considerada uma relação causal.

Todavia, há muitos outros aspectos a considerar como parte da relação causal da mortalidade infantil. Uma delas esta relacionada a própria imprecisão que existe para obtê-la e que pode vir a ser um problema permanente tornando ainda mais difícil a prova destas possíveis relações. A lógica difusa (*fuzzy logic*), por sua vez, pode assumir um papel importante na descrição de padrões de condições da população que levam à diferentes níveis da taxa da mortalidade infantil. Esta abordagem é caracterizada pela capacidade de lidar com informação imprecisa, incerta e vaga, para gerar conhecimento preciso.

A lógica difusa tem sua aplicação em diversos segmentos [7,8]. Esta abordagem esta também presente em medicina e bioinformática com aplicações que englobam sistemas de dosagem de drogas, controle de hipertensão durante anestesia, estudos difusos de epidemias, alinhamento de seqüências, avaliação de expressão genética e até mesmo índice de obesidade [9,10,11].

Neste estudo se utilizam dois algoritmos denominados Wang & Mendel [12,13] e Gath-Geva [14,15] para obter a base de regras dos modelos difusos que descrevem como a mortalidade infantil pode ser relacionada ao produto interno bruto per capita e a razão de cobertura de imunização, por exemplo, mas sem perda de generalidade, de sarampo entre menores de um ano. Com o modelo em mãos, é possível avaliar se as relações descritas na literatura podem ser também justificadas, ou não, segundo a abordagem proposta neste trabalho. Pretende-se ainda avaliar como diferentes mecanismos de aprendizado para o modelo difuso apresentam sistemas resultantes compatíveis para serem utilizados na prática.

Materiais e Método

Os dados de mortalidade infantil (Mort_inf), produto interno bruto per capita (PPP) e a razão de cobertura urbano-rural de imunização de sarampo entre menores de um ano (CIUR) foi fornecido pelo sistema de informação estatístico da Organização Mundial da Saúde (OMS) [16]. Os dados se referem ao ultimo período disponível que varia de 2002 a 2005 e englobam todos os países que fazem parte desta organização em maio de 2008.

Este artigo tem como objetivo primeiro o uso do raciocínio aproximado implementado

através de uma base de regras geradas com conjuntos e operadores difusos para mostrar as relações da mortalidade infantil com o produto interno bruto e cobertura de sarampo.

Para a extração da base de regras e das funções de pertinência que compõem a base de conhecimento foram utilizados o algoritmo Wang & Mendel, cujo funcionamento é baseado em estrutura, e o algoritmo Gath-Geva, que é baseado em agrupamento fixo. Os parâmetros que compõem os universos de discurso de entrada são a cobertura de imunização e produto interno bruto enquanto o universo de discurso de saída é dado pela taxa de mortalidade infantil. Para a atividade de extração das regras, método de Wang & Mendel é configurado para cinco funções de pertinência, enquanto, de maneira similar, o método de Gath-Geva é configurado para cinco agrupamentos (*clusters*). Ambos os métodos são ajustados para empregar funções de pertinência gaussianas no domínio de entrada; função de pertinência de saída é selecionada para ser singular (*singleton*); e o operador de conjunção escolhido foi o mínimo.

Os modelos serão utilizados para identificar algumas das relações já descritas para a mortalidade infantil na literatura e apresentar padrões (relações) pouco sustentados pelos modelos de regressão multivariada.

Resultados

Para a extração das regras e funções de pertinência foi utilizada uma amostra de 64 países com dados para os três universos de discurso. Essa amostra permitiu gerar dois modelos empregando os dois métodos Wang & Mendel (WM) e Gath-Geva (GG) mencionados anteriormente. A base de regras para o modelo difuso gerado pelo método WM é mostrado na Tabela 1. A base de regras para o modelo difuso gerado pelo método GG é mostrado na Tabela 2.

O particionamento dos universo de discurso de entrada e saída no modelo difuso pelo método WM podem ser vistos na Fig. 1. Os universos de discurso de entrada são particionados por funções de pertinência gaussianas e apresentam cinco partições equidistantes. Já o modelo difuso obtido pelo método GG pode ser visualizado na Fig. 2. Os universos de discurso de entrada são também particionados por funções de pertinência gaussianas e apresentam três partições distintas umas das outras.

Nos dois modelos difusos tem-se que o universo de discurso CIUR é particionado no intervalo de [0,9 2,4] onde o valor menor refere-se a maior percentagem de cobertura na região rural e o valor maior refere-se à maior percentagem de cobertura na região urbana, e o universo de discurso PPP é particionado entre [640 a 13170] dólares estadunidenses.

Tabela 1: Base de regras método WM

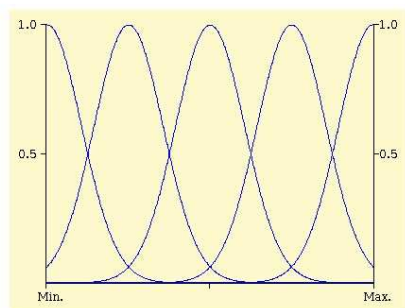
		PPP				
		mf0	mf1	mf2	mf3	mf4
CIUR	mf0	mf6	mf8	mf7	mf2	
	mf1	mf1	mf0	mf3	mf11	mf13
	mf2	mf4	mf10	mf9		
	mf3	mf5		mf14		
	mf4	mf12				

mf: Função de pertinência de um universo de discurso

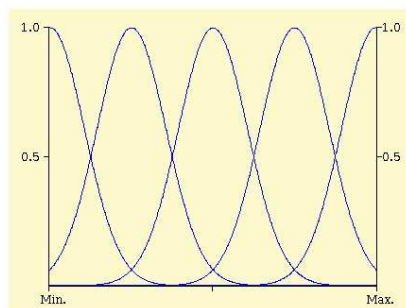
Tabela 2: Base de regras método GG

		PPP				
		mf0	mf1	mf2	mf3	mf4
CIUR	mf0	mf0				
	mf1		mf1			
	mf2			mf2		
	mf3				mf3	
	mf4					mf4

mf: Função de pertinência de um universo de discurso

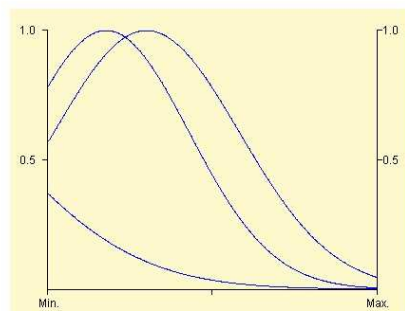


CIUR

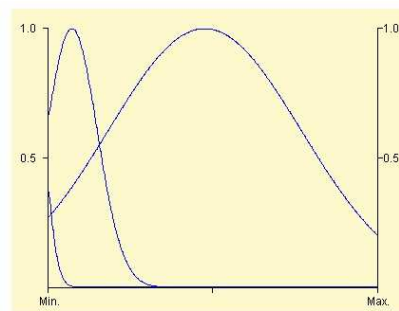


PPP

Figura 1 – Partição dos universo de discurso CIUR e PPP no modelo de WM.



CIUR

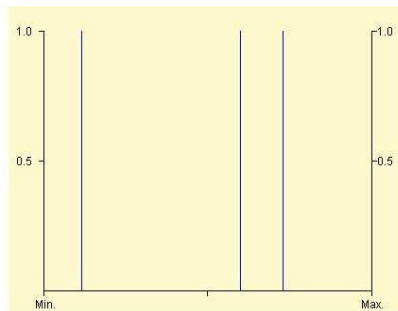


PPP

Figura 2 – Partição dos universo de discurso CIUR e PPP no modelo de GG.



Wang & Mendel



Gath-Geva

Figura 3 – Partição dos universo de discurso Mortalidade Infantil dos modelos de WM e GG.

Modelo gerado pelo método WM

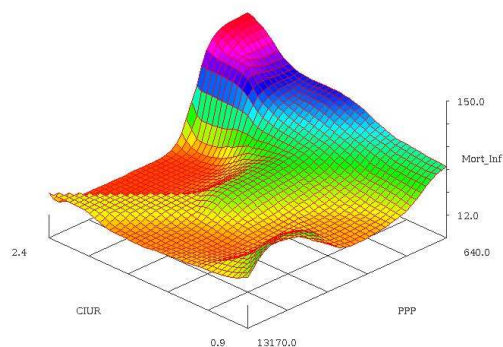


Figura 4 – Modelo difuso para a Mortalidade Infantil por WM.

Modelo gerado pelo método GG

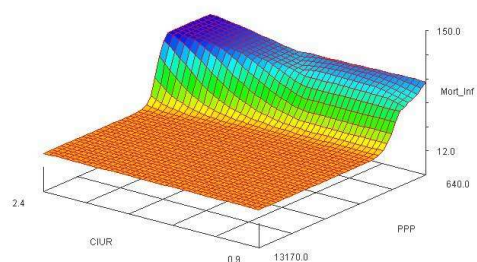


Figura 5 – Modelo difuso para a Mortalidade Infantil por GG.

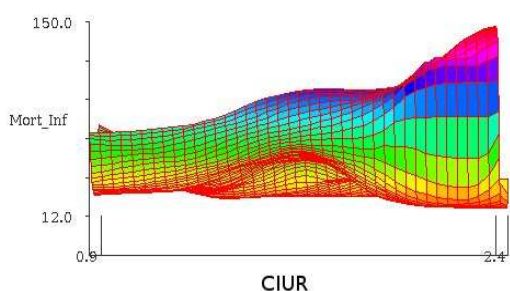


Figura 6 – Vista 2D do universo de discurso CIUR no modelo de WM .

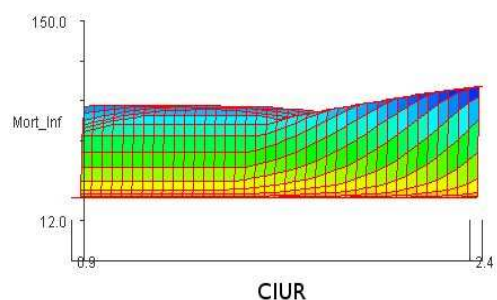


Figura 8 - Vista 2D do universo de discurso CIUR no modelo de GG.

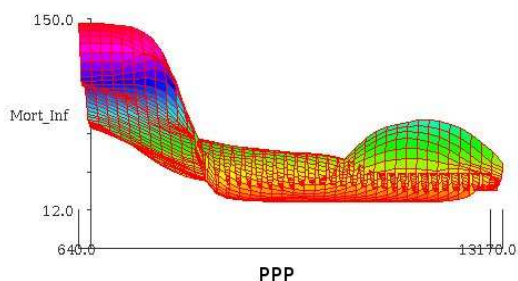


Figura 7 - Vista 2D do universo de discurso PPP no modelo de WM.

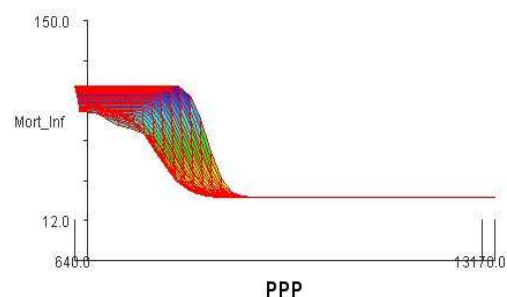


Figura 9 - Vista 2D do universo de discurso PPP no modelo de GG.

A mortalidade infantil foi particionada entre os valores [12 a 250] por 1000 nascidos vivos e é mostrada na Fig. 3 para os dois modelos.

As bases de conhecimento resultantes em gráficos 3D são apresentadas na Fig. 4 e 5, respectivamente, para o modelo difuso para a mortalidade infantil por WM e modelo difuso para a mortalidade infantil por GG.

As projeções do primeiro modelo difuso no universo de discurso CIUR e PPP, respectivamente, são vistas nas Fig. 6 e 7. As

projeções do modelo difuso obtido pelo método GG são apresentadas nas Fig. 8 e 9.

Discussão e Conclusões

Uma clara diferença é observada entre os modelos de WM e GG. Esta diferença se deve à configuração dos universos de entrada e saída (Fig 1, 2 e 3) e a base de conhecimento gerada (Fig. 4 e 5). Estas diferenças são explicadas porque que WM é um algoritmo que requer pre-

configuração por parte do programador no universo de discurso de entrada; as funções de pertinência assumem uma distribuição fixa e porque ele é feito para dar expressividade à representação do conhecimento inserido na base de dados e, como tal, não há limite no número de regras [11]. Por outro lado, o método GG é um algoritmo que é limitado pelo programador no número máximo de agrupamentos (clusters) a serem modelados dentro da base de dados [13] e, por isto, suas funções de pertinência assumem uma distribuição adaptativa e móvel.

Quando as bases de conhecimento dos modelos geradas pelos métodos WM e GG são comparadas observa-se como a pre-configuração afeta a representação da base de dados. São afetados o número de regras e quantidade de funções de pertinência, como apresentadas no interior das Tabelas 1 e 2. Estas diferenças se refletem na forma das superfícies de mapeamento entre entrada-saída, como mostrado na Fig. 4 e 5.

Se bem existem diferenças visuais entre as bases de conhecimento, a projeção nos universos de discurso, (Fig. 6, 7 e Fig. 8, 9) permitem perceber que as diferenças descrevem um padrão com certa similaridade em termos de distribuição. Ao comparar as Fig. 6 e 8 é possível verificar que a mortalidade infantil tem uma faixa que é correspondente com todo o intervalo de valores CIUR para ambos os modelos. No caso das Fig. 7 e 9 a relação é muito mais evidente onde a mortalidade infantil parece decrescer quando o PPP aumenta.

Os resultados são similares em relação ao ingresso nacional bruto per capita como achados na literatura [5, 6, 17]. Um aspecto importante que permitem evidenciar os modelos é a mortalidade infantil não é descrita só pelo PPP.

Neste trabalho foi usada a razão de cobertura urbano-rural da imunização entre menores de um ano. Esta variável representa um importante aspecto da saúde da população, como o são os programas de vacinação (Sarampo, DTP3, HepB3, entre outros) para o controle da mortalidade infantil [18, 19, 20].

O modelo de GG resulta em uma projeção da relação da mortalidade infantil e PPP estendida para todo o intervalo de CIUR. Uma importante conclusão é obtida. Esta relação se faz mais ampla na medida que CIUR se incrementa, com isso verificamos que existem países com baixo PPP que apresentam uma taxa de mortalidade infantil baixa. No caso do modelo de WM inclui as relações de GG, mas define uma área onde um alto PPP com moderada CIUR apresentam uma taxa de mortalidade infantil mais elevada. Se é considerado que a cobertura de um serviço de saúde significa dar acesso a um serviço por meio da destinação de recurso do estado, então pode-se pensar nela como uma forma de redistribuição. Desta forma, os modelos

permitem extrapolar que existe um efeito de redistribuição que explica o nível da mortalidade infantil conforme defendido em [6,17], mas não apresenta uma medida da força da relação como o proposto em [4].

Trabalhos futuros devem explorar mais relações da mortalidade infantil com outras variáveis que descrevem a saúde da população aproveitando as bases de dados disponíveis na OMS, entre outras. Pretende-se com isto, produzir sistemas de decisão baseados em conhecimento que permitam fazer a gestão do monitoramento da mortalidade infantil visto que ela é, em muitas situações, evitável. A vantagem do sistema de decisão e, em particular, aqueles baseados em lógica difusa, é que se tem um mecanismo para apoiar gestores na tomada de decisões com respeito aos programas de saúde, assim como identificar possíveis causas que permitam prevenir a mortalidade infantil localmente e medir o impacto dessas decisões sobre a qualidade de vida da população. Como mostrado neste artigo, é possível que isto seja alcançado ao se aplicar os princípios que a Inteligência Computacional dispõe através da abrangente gama de ferramentas atualmente disponíveis. Neste sentido a lógica difusa se torna uma ferramenta poderosa, pois permite lidar com incertezas, imprecisões, vaguidade e verdades parciais, um tipo de informação presente em biomedicina e saúde. Quando associada a ferramentas de aprendizado e otimização como, p.ex., os métodos Wang & Mendel (WM) e Gath-Geva (GG), então não só é possível lidar com informações imperfeitas, mas também a aprender com elas.

Por fim, um breve resumo do trabalho observa-se que o uso da lógica difusa mostrou-se como uma alternativa para avaliar aspectos da mortalidade que são controversos. Esta abordagem vem a ser uma ferramenta efetiva quando se utilizam dados muitas vezes imprecisos e incertos de diversas fontes para explorar o significado da mortalidade infantil.

Referências

1. Pampel FC, Jr., Pillai VK. Patterns and determinants of infant mortality in developed nations, 1950-1975. *Demography*. 1986 Nov;23(4):525-42.
2. Murray CJ, Chen LC. In search of a contemporary theory for understanding mortality change. *Soc Sci Med*. 1993 Jan;36(2):143-55.
3. Bulterys M. In search of a contemporary theory for understanding mortality change. *Soc Sci Med*. 1995 Apr;40(8):1165.
4. Judge K, Mulligan JA, Benzeval M. Income inequality and population health. *Soc Sci Med*.

1998 Feb-Mar;46(4-5):567-79.

5. Hales S, Howden-Chapman P, Salmond C, Woodward A, Mackenbach J. National infant mortality rates in relation to gross national product and distribution of income. *Lancet*. 1999 Dec 11;354(9195):2047.

6. Babones SJ. Income inequality and population health: correlation and causality. *Soc Sci Med*. 2008 Apr;66(7):1614-26.

7. Zadeh L. Fuzzy Logic. *Computer*. 1988;21(4):83-93.

8. Lee CC. Fuzzy logic in control systems: fuzzy logic controller. II. Systems, Man and Cybernetics, *IEEE Transactions on*. 1990;20(2):419-35.

9. Torres A, Nieto JJ. Fuzzy logic in medicine and bioinformatics. *J Biomed Biotechnol*. 2006;2006(2):91908.

10. Mahfouf M, Abbod MF, Linkens DA. A survey of fuzzy logic monitoring and control utilisation in medicine. *Artif Intell Med*. 2001 Jan-Mar;21(1-3):27-42.

11. Miyahira, S. Araujo, E., "Fuzzy Obesity Index for Obesity Treatment and Surgical Indication", *Proc. IEEE World Congress on Computational Intelligence (WCCI 2008) / IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, (pre-print), Hong Kong, 2008.

12. Li-Xin W. The WM method completed: a flexible fuzzy system approach to data mining. *Fuzzy Systems, IEEE Transactions on*. 2003;11(6):768-82.

13. Casillas J, Cord O, Herrera F. Improving the wang and mendel's fuzzy rule learning method by inducing cooperation among rules. *Proceedings of the 8th Information Processing and Management of Uncertainty in Knowledge-Based Systems Conference*. 2000:1682-8.

14. Hasanzadeh RPR., Moradi MH, SHH S. Fuzzy Clustering to the Detection of Defects from Nondestructive Testing. 3rd International Conference: Sciences of electronic Technologies of Information and Telecommunications; 2005; Tunisia..

15. Gil-Garcia RJ, Badía JM. Algoritmos de Agrupamiento. Castellón: Departamento de Ingeniería y Ciencia de Computadores / Universidad Jaime I; 2002 Contract No.: Document Number].

16. WHO. WHO Statistical Information System (WHOSIS). 2008 [updated 2008; cited]; Available from: <http://www.who.int/whosis/en/>.

17. Wilkinson RG, Pickett KE. Income inequality and population health: a review and explanation of the evidence. *Soc Sci Med*. 2006 Apr;62(7):1768-84.

18. Quadros C. Un siglo de vacunas e inmunización en las Américas. In: PAHO, editor. *Vacunas: Prevención de enfermedades y protección de la salud*. Washington: PAHO Publications; 2004. p. 452.

19. Davey S. Vacunas e Inmunización: situación mundial. OMS; 2003 [cited. Available from: http://www.who.int/vaccines-documents/DocsPDF05/Vacunas_e_inmunizacion_completo.pdf.

20. Jacobson Vann JC, Szilagyi P. Patient reminder and recall systems to improve immunization rates. *Journal* [serial on the Internet]. 2005 Date; (3): Available from: <http://www.mrw.interscience.wiley.com/cochrane/cisysrev/articles/CD003941/frame.html>.

Contato

Milton Urrutia, milton-pg@dis.epm.br.

Ernesto Araujo, Prof. Dr. ernesto.araujo@unifesp.br