

Modelos de predição utilizando lógica fuzzy: uma abordagem inspirada na inferência Bayesiana

Felipo Bacani¹ Laécio C. de Barros²
DMA, IMECC, UNICAMP – 13.083-970, Campinas/SP.

Resumo. O presente trabalho tem por objetivo aplicar a teoria de conjuntos fuzzy a modelos de predição (inferência) de dados. O modelo utilizado baseia-se em relações fuzzy em espaços contínuos, e a analogia com a inferência bayesiana é feita reinterpretando implicações fuzzy que aparecem no processo. Entretanto, são apenas analogias conceituais: o presente trabalho não lida com nenhuma distribuição de probabilidades. A metodologia proposta é utilizada com o objetivo de tornar mais precisa a previsão de um especialista, levando em conta um registro histórico sobre o problema, ou seja, melhorar a previsão do especialista levando em conta o que ocorreu com as previsões anteriores.

Palavras-chave: *Conjuntos fuzzy; metodologia bayesiana; modelos de predição; Biomatemática.*

1. Introdução

O presente trabalho utiliza o método proposto em Lapointe e Bobée (2000), que tem por objetivo melhorar a previsão fornecida por um especialista utilizando um registro histórico, de forma a também levar em conta *o que ocorreu com as previsões anteriores* (Bacani, 2012).

Dessa forma, tem-se como objetivo propor um modelo matemático que “capte” o conhecimento do especialista e, se possível, melhore as previsões feitas por ele. Isso será feito utilizando um registro histórico que contém informações

¹bacani@ime.unicamp.br

²laeciocb@ime.unicamp.br

sobre as previsões anteriores do especialista e as medidas correspondentes. Designando:

- X : Variável que se pretende prever (temperatura, por exemplo)
- Y : Previsão da variável X , fornecida por um especialista,

pretende-se melhorar a previsão Y sobre a variável X , levando em conta o registro histórico.

Usualmente em tais problemas utiliza-se um modelo equivalente obtido através da teoria de inferência bayesiana. Algumas das vantagens metodológicas do método proposto, se comparado ao modelo equivalente bayesiano:

- Uma solução analítica pode sempre ser obtida no caso fuzzy, conforme o teorema 1. Isso nem sempre ocorre no caso bayesiano, pois muitas vezes é preciso realizar aproximações numéricas de integrais, o que não ocorre no modelo fuzzy;
- Ao lidar com poucos dados, o método bayesiano requer modelos demasiadamente específicos cujos parâmetros precisam ser precisamente estimados. Por outro lado, na abordagem fuzzy, pouca informação, muitas vezes, é suficiente para obter modelos satisfatórios.

Os primeiros trabalhos que foram na direção do artigo de (Lapointe e Bobée, 2000) datam dos anos 70, a partir de questionamentos do próprio Zadeh (criador da teoria de conjuntos fuzzy) que foram levados adiante por (Nguyen, 1978) e (Hisdal, 1978), estabelecendo sempre uma relação com a teoria de probabilidades. Talvez o primeiro trabalho a observar o problema como a resolução de equações relacionais fuzzy foi feito por (Bouchon, 1987), que (assim como (Lapointe e Bobée, 2000)) observa uma implicação fuzzy como uma relação *condicional* entre o antecedente e o conseqüente do *modus ponens*.

1.1. Fundamentos Teóricos

1.1.1 Inferência Bayesiana

A realização de previsão utilizando a metodologia bayesiana é frequentemente utilizada em problemas em que há informação histórica escassa ou inútil no momento que a previsão inicial é necessária.

Seja y uma variável aleatória com função de densidade f , caracterizada por um parâmetro desconhecido x . Escreve-se essa densidade como $f(y|x)$

para indicar que a distribuição de probabilidades depende do valor de x . O parâmetro x é uma variável aleatória com função de densidade $h_0(x)$, que é chamada a densidade *priori* de x . Esta distribuição de probabilidades mede a informação subjetiva (ou “grau de confiança”) sobre o valor de x .

Em uma situação de previsão, suponha que a estimativa inicial de x é dada como uma distribuição de probabilidades $h_0(x)$, e que a informação subsequente sobre o fenômeno é obtida na forma de uma variável aleatória y , cuja distribuição de probabilidades $f(y|x)$ (chamada de *função de verossimilhança*) depende de x . A nova estimativa de x será na forma de uma distribuição revisada, $h_1(x|y)$, chamada distribuição *posteriori*. Se as variáveis aleatórias x e y forem contínuas, pelo teorema de Bayes (Montgomery, 1976, p. 241):

$$h_1(x|y) = \frac{f(y|x) h_0(x)}{\int_x f(y|x) h_0(x) dx} = \frac{f(y|x) h_0(x)}{g(y)}. \quad (1)$$

A distribuição posteriori $h_1(x|y)$ pode ser vista como um “acordo formal” entre a verossimilhança $f(y|x)$, que resume a evidência nos dados por si só, e a distribuição priori $h_0(x)$, que resume as evidências externas que sugerem frequências maiores.

1.1.2 Distribuições de possibilidades

Neste trabalho escolhe-se por representar conjuntos fuzzy como distribuições de possibilidade, pelo fato de ser mais diretamente comparada à teoria de probabilidades que conjuntos fuzzy. Isso se deve ao fato da teoria de possibilidades propor (como probabilidades) funções associadas a conjuntos que quantificam a incerteza de eventos (Dubois et al., 1997).

Pode-se enxergar uma distribuição de possibilidades como uma representação da função de pertinência φ_X de um conjunto fuzzy X (Zadeh, 1978). A distribuição de possibilidades de uma variável X , que toma valores em um conjunto universo U , é portanto, uma função de U em $[0, 1]$, denotada por $\pi_X(\cdot)$.

1.1.3 Relações fuzzy

Definição 1.1 (Relação fuzzy) *Uma relação fuzzy \mathcal{R} sobre $U_1 \times U_2 \times \dots \times U_n$ é qualquer subconjunto fuzzy de $U_1 \times U_2 \times \dots \times U_n$. Assim, uma relação fuzzy \mathcal{R} é definida por uma função de pertinência $\varphi_{\mathcal{R}} : U_1 \times U_2 \times \dots \times U_n \rightarrow [0, 1]$.*

Definição 1.2 (Regra da composição) *Sejam U e V dois conjuntos, $\mathcal{F}(U)$ e $\mathcal{F}(V)$ as classes de subconjuntos fuzzy de U e V respectivamente, e \mathcal{R} uma relação fuzzy sobre $U \times V$. A relação \mathcal{R} define um funcional de $\mathcal{F}(U)$ em $\mathcal{F}(V)$ que, a cada elemento $A \in \mathcal{F}(U)$, faz corresponder o elemento $B \in \mathcal{F}(V)$ cuja função de pertinência é dada por*

$$\varphi_B(y) = \sup_{x \in U} \{ \varphi_{\mathcal{R}}^T(x, y) \wedge \varphi_A(x) \} = \sup_{x \in U} \{ \min(\varphi_{\mathcal{R}}(x, y), \varphi_A(x)) \}. \quad (2)$$

Esta expressão é conhecida como regra da composição.

Trocando o operador \wedge por uma t-norma $T = \Delta$ arbitrária, a regra de composição (Def. 1.2) toma a forma:

$$\varphi_B(y) = \sup_{x \in U} \{ \varphi_{\mathcal{R}}^T(x, y) \Delta \varphi_A(x) \} = \sup_{x \in U} \{ T(\varphi_{\mathcal{R}}^T(x, y), \varphi_A(x)) \}, \quad (3)$$

onde $\varphi_{\mathcal{R}}^T$ é a função de pertinência de uma *implicação fuzzy*, que é uma regra que leva valores de pertinência do conjunto A em valores de pertinência de B . Pode-se denotar, portanto, esta implicação por $(\varphi_A \implies \varphi_B)$.

Existem algumas complicações na obtenção de $\varphi_{\mathcal{R}}$, pois os mesmos conjuntos A e B resultam, em geral, duas relações fuzzy distintas, caso se utilize duas t-normas distintas. Isso faz com que ela seja denotada por $\varphi_{\mathcal{R}}^T$, em função da t-norma T .

Supondo que os conjuntos A e B sejam dados, é possível que exista mais de uma solução $\varphi_{\mathcal{R}}^T$ para a equação (3), e, como critério de escolha, busca-se comumente a solução maximal (a “maior” solução).

Definição 1.3 (Solução maximal) *Seja $\overline{\varphi}_{\mathcal{R}}^T$ uma relação solução de um problema relacional fuzzy. Se ocorrer que*

$$\forall x \in U, y \in V, \quad \overline{\varphi}_{\mathcal{R}}^T(x, y) \geq \varphi_{\mathcal{R}}^T(x, y),$$

para toda solução $\varphi_{\mathcal{R}}^T$, então $\overline{\varphi}_{\mathcal{R}}^T$ é a solução maximal do problema em questão.

Um dos principais resultados teóricos deste trabalho é o Teorema 1, enunciado abaixo (ver (Barros e Bassanezi, 2010, p.176)). É ele que permite obter a solução $\varphi_{\mathcal{R}}^T(x, y)$ da equação 3 de forma analítica.

Teorema 1 (Existência da solução) *Seja a equação relacional fuzzy (3):*

$$\varphi_B(y) = \sup_{x \in U} \{ T(\varphi_{\mathcal{R}}^T(x, y), \varphi_A(x)) \}$$

Dada uma *t*-norma T e os conjuntos fuzzy $A \in \mathcal{F}(U)$ e $B \in \mathcal{F}(V)$, a relação \mathcal{R} em $(U \times V)$, que é solução maximal¹ do problema acima, tem função de pertinência dada pela expressão:

$$\varphi_{\mathcal{R}}^T(x, y) = \sup_z \{z \in [0, 1] : T(z, \varphi_A(x)) \leq \varphi_B(y)\} = \psi_T(\varphi_A(x), \varphi_B(y)),$$

onde $\psi_T(a, b) = \sup_z \{z \in [0, 1] : T(z, a) \leq b\}$

Demonstração. A prova é encontrada em (Pedrycz, 1982).

Neste trabalho foram utilizadas duas *t*-normas: a do mínimo (\wedge) e a do produto (\cdot). Para estes dois casos particulares de *t*-norma, o teorema 1 toma a seguinte forma:

Corolário 1 (Solução para *t*-norma do produto) Se $T = \text{produto}$,

$$\psi_{(\cdot)}(\varphi_A(x), \varphi_B(y)) = \begin{cases} 1 & \text{se } \varphi_A(x) < \varphi_B(y) \\ \frac{\varphi_B(y)}{\varphi_A(x)} & \text{se } \varphi_A(x) \geq \varphi_B(y) \end{cases}$$

Corolário 2 (Solução para *t*-norma do mínimo) Se $T = \text{mínimo}$,

$$\psi_{(\wedge)}(\varphi_A(x), \varphi_B(y)) = \begin{cases} 1 & \text{se } \varphi_A(x) \leq \varphi_B(y) \\ \varphi_B(y) & \text{se } \varphi_A(x) > \varphi_B(y) \end{cases}$$

2. Metodologia

2.1. Regra de composição/Analogia com a inferência bayesiana

No método proposto, pretende-se obter informações sobre a variável X com base em Y , uma previsão de X . Ao observar este problema a partir de relações fuzzy (eq. 3), que obtém informações sobre o conjunto B com base em informações sobre A através da implicação ($\varphi_A \implies \varphi_B$), no problema em questão esta implicação torna-se ($Y \implies X$).

Representando X e Y como as distribuições de possibilidade $\pi_X \equiv \varphi_B$ e $\pi_Y \equiv \varphi_A$, a implicação torna-se " $\pi_Y \implies \pi_X$ ". Como implicações são proposições lógicas *condicionais* do tipo "Se p então q ", é possível denotar a função de

¹Se num problema de relações fuzzy houver uma única solução, esta solução é a maximal. Portanto, se não houver solução maximal possível, não existe solução para o problema em questão.

pertinência da relação fuzzy $\varphi_{\mathcal{R}}^T$ como a *distribuição condicional de possibilidades* $\pi_{X|Y}^T$, pois representa o conhecimento sobre a variável X dada condicionada à informação sobre a variável Y . Reescrevendo a fórmula (3) em função das distribuições de possibilidades, tem-se:

$$\pi_X^T(x) = \sup_y \left\{ \pi_{X|Y}^T(x|y) \Delta \pi_Y(y) \right\} = \sup_y \left\{ T \left(\pi_{X|Y}^T(x|y), \pi_Y(y) \right) \right\}. \quad (4)$$

Observando a equação da distribuição conjunta de probabilidades,

$$\mathbf{P}_{X,Y}(x, y) = \mathbf{P}_{X|Y}(x|y) \cdot \mathbf{P}_Y(y),$$

e comparando seu lado direito com o termo “ $\pi_{X|Y}(x|y) \Delta \pi_Y(y)$ ” da equação (4), define-se

$$\pi_{X,Y}^T(x, y) = \pi_{X|Y}^T(x|y) \Delta \pi_Y(y) = T \left(\pi_{X|Y}^T(x|y), \pi_Y(y) \right) \quad (5)$$

como a *distribuição conjunta de possibilidades* de X, Y . De forma análoga, pode-se também definir

$$\pi_{X,Y}^T(x, y) = \pi_{Y|X}^T(y|x) \Delta \pi_X(x) = T \left(\pi_{Y|X}^T(y|x), \pi_X(x) \right). \quad (6)$$

é importante notar que as distribuições probabilísticas foram trocadas pelas análogas possibilísticas (fuzzy) nas equações (5) e (6), e o operador de produto foi trocado por uma t-norma T arbitrária.

Agora, observando a equação da distribuição marginal de possibilidades da variável aleatória Y :

$$\mathbf{P}_Y(y) = \int_x \mathbf{P}(y|x) \cdot \mathbf{P}_X(x) dx,$$

e substituindo o operador de integral em y por supremo em y e comparando com a equação (4), pode-se designar

$$\pi_Y^T(y) = \sup_x \left\{ T \left(\pi_{Y|X}^T(y|x), \pi_X(x) \right) \right\} = \sup_x \left\{ \pi_{X,Y}^T(x, y) \right\}, \quad (7)$$

como a *distribuição marginal de possibilidades* de Y . $\pi_{Y|X}^T(y|x)$ é a distribuição condicional de possibilidades de Y dado X , segundo a t-norma $T = \Delta$.

As interpretações da função de verossimilhança e das distribuições priori e posteriori têm neste trabalho o mesmo significado que no caso bayesiano (Seção 1.1.1).

$$\begin{array}{ccc}
 \left. \begin{array}{l} \text{(priori)} \ h_0(x) \\ \text{(verossimilhança)} \ f(y|x) \end{array} \right\} h_1(x|y) \text{ (posteriori)} & & \left. \begin{array}{l} \text{(priori)} \ \pi_X(x) \\ \text{(verossimilhança)} \ \pi_{Y|X}(y|x) \end{array} \right\} \pi_{X|Y}^T(x|y) \text{ (posteriori)}
 \end{array}$$

(a) Processo de inferência bayesiano. (b) Processo proposto.

Figura 1: Analogia entre os processos de inferência bayesiano e o proposto.

Dessa forma, os dois padrões de inferência são análogos, conforme a figura 1.

Nota-se também que na presente abordagem a priori $\pi(x)$ e a função de verossimilhança $\pi(y|x)$ *não dependem da t-norma T*, pois suas expressões são obtidas a partir do registro histórico em mãos. Para um conjunto de dados de exemplo, explica-se como obtê-las na seção 3.1.

2.2. Obtendo a distribuição posteriori

Como será visto adiante, existem duas formas possíveis de se obter a distribuição posteriori $\pi^T(x|y)$. Com base em alguns argumentos, apenas uma das duas formas será escolhida.

2.2.1 Forma 1: Utilizando a Regra de Composição

Lembrando que a equação (4),

$$\text{Condição 1: } \pi(x) = \sup_y \{T(\pi^T(x|y), \pi^T(y))\}, \quad (8)$$

é apenas uma reinterpretação da equação relacional (3), pode-se, utilizando o teorema 1, determinar $\pi^T(x|y)$ em função de $\pi(x)$ e $\pi^T(y)$.

Assim, é possível obter a primeira fórmula para a distribuição posteriori, denotada por $\pi_1^T(x|y)$:

$$\pi_1^T(x|y) = \psi_T(\pi^T(y), \pi(x)). \quad (9)$$

2.2.2 Forma 2: Forçando a consistência da distribuição conjunta

Procura-se por soluções em que haja consistência entre as duas formas de definir a distribuição conjunta, ou seja, que as eqs. (5) e (6) sejam iguais:

$$\text{Condição 2: } \forall x, y, \quad T(\pi^T(x|y), \pi^T(y)) = T(\pi(y|x), \pi(x)) \quad (10)$$

As distribuições posteriori $\pi^T(x|y)$ que satisfazem a condição 2 (eq. 10) acima, são os elementos “ z ” do conjunto

$$\{z \in [0, 1] : T(z, \pi^T(y)) = T(\pi(y|x), \pi(x))\}. \quad (11)$$

Como se procura apenas pela “maior” posteriori (ou solução maximal, no contexto de relações fuzzy), toma-se o supremo entre as soluções do conjunto definido na eq. 11 acima (ou seja, toma-se supremo em z), de forma a obter \bar{z} :

$$\bar{z} = \sup_z \{z \in [0, 1] : T(z, \pi^T(y)) = T(\pi(y|x), \pi(x))\},$$

e, pela monotonicidade da t-norma T (Lapointe e Bobée, 2000, p. 128),

$$\bar{z} = \sup_z \{z \in [0, 1] : T(z, \pi^T(y)) \leq T(\pi(y|x), \pi(x))\}. \quad (12)$$

Representando o conjunto da equação 12 acima através da notação do teorema 1:

$$\bar{z} = \psi_T(\pi^T(y), T(\pi(y|x), \pi(x))). \quad (13)$$

A equação 13 permite obter a “maior solução” \bar{z} utilizando o mesmo teorema 1 que foi utilizado na “forma 1” (seção 2.2.1), pois ambas estão definidas utilizando o conjunto ψ_T .

Assim, a eq. 13 é a segunda forma de se obter a distribuição posteriori:

$$\pi_2^T(x|y) = \psi_T(\pi^T(y), T(\pi(y|x), \pi(x))). \quad (14)$$

2.2.3 Escolhendo entre as duas soluções

Duas formas de obter a distribuição posteriori $\pi(x|y)$ foram obtidas, respectivamente $\pi_1^T(x|y)$ (9) e $\pi_2^T(x|y)$ (14), ambas em função do conjunto ψ_T (teorema 1).

A escolha entre as soluções $\pi_1^T(x|y)$ e $\pi_2^T(x|y)$ leva dois fatos em consideração:

1. Conforme a observação abaixo, nem toda posteriori obtida através da “Forma 1”, ou seja, que satisfaça a condição 1 (eq. 8) satisfaz a condição 2 (eq. 10);
2. Por outro lado, mostra-se (teorema 2) que *toda* posteriori obtida pela “Forma 2”, ou seja, que satisfaça a condição 2 (eq. 10) satisfaz também a condição 1 (eq. 8).

Observação: Uma posteriori $\pi_1^T(x|y)$ dada pela eq. 9, solução maximal que satisfaz a condição 1 (eq. 8), em geral, não satisfaz a condição 2 (eq. 10). Mostra-se este resultado através de um contra-exemplo:

Suponha a t-norma do produto, e a posteriori $\hat{\pi}_1^{(\cdot)}(x|y)$ dada pela eq. 9, que por construção satisfaz a condição 1. Seja $\hat{\pi}_1^{(\cdot)}(x|y)$ tal que:

1. $\forall x, y, \pi(x) \geq \pi^{(\cdot)}(y) > 0$
2. Existe \bar{x} tal que, $\forall y, \pi(\bar{x}) > \pi^{(\cdot)}(\bar{x}, y)$

Aplicando o corolário 1 à equação 9 (e usando o item 1 acima), obtém-se uma distribuição posteriori contínua:

$$\hat{\pi}_1^{(\cdot)}(x|y) = \frac{\pi(x)}{\pi^{(\cdot)}(y)}.$$

Resta saber se a distribuição posteriori $\hat{\pi}_1^{(\cdot)}(x|y)$ satisfaz a condição 2 (eq. 10). Aplicando $\hat{\pi}_1^{(\cdot)}(x|y)$ no ponto \bar{x} no lado esquerdo da condição 2,

$$\begin{aligned} T\left(\hat{\pi}_1^{(\cdot)}(\bar{x}|y), \pi^{(\cdot)}(y)\right) &= \hat{\pi}_1^{(\cdot)}(\bar{x}|y) \cdot \pi^{(\cdot)}(y) \\ T\left(\hat{\pi}_1^{(\cdot)}(\bar{x}|y), \pi^{(\cdot)}(y)\right) &= \left(\frac{\pi(\bar{x})}{\pi^{(\cdot)}(y)}\right) \cdot \pi^{(\cdot)}(y) = \pi(\bar{x}) > \pi^{(\cdot)}(\bar{x}, y). \end{aligned}$$

Então $T\left(\hat{\pi}_1^{(\cdot)}(x|y), \pi^T(y)\right) \neq \pi^{(\cdot)}(x, y)$ para algum x , ou seja, a condição 2 (eq. 10) não foi satisfeita. Dessa forma, foi obtida uma posteriori $\hat{\pi}_1^{(\cdot)}(x|y)$ que satisfaz a condição 1 (eq. 8) mas que não satisfaz a condição 2 (eq. 10).

Teorema 2 Uma posteriori $\pi_2^T(x|y)$ dada pela equação (14), solução maximal que satisfaz a condição 2 (eq. 10), sempre satisfaz a condição 1 (eq. 8).

Prova: 1 Demonstração: Suponha $\pi_2^T(x|y)$ uma distribuição posteriori dada pela eq. 14. Por construção ela satisfaz a condição 2 (eq. 10):

$$T\left(\pi_2^T(x|y), \pi^T(y)\right) = T\left(\pi(y|x), \pi(x)\right) = \pi^T(x, y)$$

Tomando supremo em y e substituindo $\pi(x) = \sup_y \{\pi^T(x, y)\}$ (eq. 7),

$$\begin{aligned} \sup_y \{T\left(\pi_2^T(x|y), \pi^T(y)\right)\} &= \sup_y \{\pi^T(x, y)\} \\ \sup_y \{T\left(\pi_2^T(x|y), \pi^T(y)\right)\} &= \pi(x), \end{aligned}$$

donde se conclui $\pi_2^T(x|y)$ satisfaz a condição 1 (eq. 8).

Escolha pela “Forma 2” de obter a posteriori:

Com base na observação acima e no teorema 2, escolhe-se obter sempre a distribuição posteriori através da “Forma 2” (eq. 14). Denota-se, então, a partir de agora $\pi_2^T(x|y) = \pi^T(x|y)$, ou seja:

$$\pi^T(x|y) = \psi_T(\pi^T(y), T(\pi(y|x), \pi(x))), \quad (15)$$

onde $\psi_T(a, b) = \sup_{z \in [0,1]} \{z : T(a, z) \leq b\}$

A seguir, são obtidas as distribuições posteriori de possibilidade acima para os casos particulares de t-norma do produto e mínimo, respectivamente, $\pi^{(\cdot)}(x|y)$ (corolário 1) e $\pi^{(\wedge)}(x|y)$ (corolário 2).

Posteriori $\pi^{(\cdot)}(x|y)$, t-norma=produto:

Aplicando o corolário 1 para $\varphi_A(x) = \pi^{(\cdot)}(y)$ e $\varphi_B(y) = \pi(y|x) \cdot \pi(x)$,

$$\pi^{(\cdot)}(x|y) = \begin{cases} 1 & \text{se } \pi^{(\cdot)}(y) < \pi^{(\cdot)}(x, y) \\ \frac{\pi(y|x) \cdot \pi(x)}{\pi^{(\cdot)}(y)} & \text{se } \pi^{(\cdot)}(y) \geq \pi^{(\cdot)}(x, y) \end{cases}$$

Como $\forall x, y$ tem-se que pela eq. (7), $\pi^{(\cdot)}(y) = \sup_x \{\pi^{(\cdot)}(x, y)\} \geq \pi^{(\cdot)}(x, y)$. Assim a parte onde a posteriori vale sempre “1” é descartada, logo:

$$\pi^{(\cdot)}(x|y) = \frac{\pi(y|x) \cdot \pi(x)}{\pi^{(\cdot)}(y)} \quad (16)$$

Posteriori $\pi^{(\wedge)}(x|y)$, t-norma=mínimo:

Aplicando o corolário 2 para $\varphi_A(x) = \pi^{(\wedge)}(y)$ e $\varphi_B(y) = \pi(y|x) \wedge \pi(x)$,

$$\pi^{(\wedge)}(x|y) = \begin{cases} 1 & \text{se } \pi^{(\wedge)}(y) \leq \pi^{(\wedge)}(x, y) \\ \pi(y|x) \wedge \pi(x) & \text{se } \pi^{(\wedge)}(y) > \pi^{(\wedge)}(x, y) \end{cases}$$

Da eq. (7), conclui-se que $\pi^{(\wedge)}(y) \geq \pi^{(\wedge)}(x, y)$. Dessa forma:

$$\pi^{(\wedge)}(x|y) = \begin{cases} 1 & \text{se } \pi^{(\wedge)}(y) = \pi^{(\wedge)}(x, y) \\ \pi(y|x) \wedge \pi(x) & \text{se } \pi^{(\wedge)}(y) > \pi^{(\wedge)}(x, y) \end{cases} \quad (17)$$

Comparação entre o processo de inferência proposto o bayesiano:

Conclui-se que, para o caso particular de t-norma do produto, a distribuição posteriori $\pi^{(\cdot)}(x|y)$ solução assume a forma (eq. 16):

$$\pi^{(\cdot)}(x|y) = \frac{\pi(y|x) \pi(x)}{\sup_x \{\pi(y|x) \pi(x)\}},$$

onde $\pi(x)$ é a distribuição priori e $\pi(y|x)$ a verossimilhança. Nota-se a clara semelhança entre esta expressão e a posteriori $h_1(x|y)$ do caso bayesiano (eq. 1):

$$h_1(x|y) = \frac{f(y|x) h_0(x)}{\int_x f(y|x) h_0(x) dx},$$

onde $h_0(x)$ é a priori e $f(y|x)$ a verossimilhança. Deve-se ressaltar a correspondência entre a distribuição priori ($\pi(x) \sim h_0(x)$) e verossimilhança ($\pi(y|x) \sim f(y|x)$) nos dois casos, o que permite notar mais claramente a analogia das duas metodologias, conforme o ilustra a figura 1.

2.3. Resumo das equações da metodologia

Com o objetivo de facilitar a consulta, as equações utilizadas na metodologia são reproduzidas a seguir.

$$\pi^T(x, y) = T(\pi(y|x), \pi(x)) \quad (18)$$

$$\pi^T(y) = \sup_x \{\pi^T(x, y)\} = \sup_x \{T(\pi(y|x), \pi(x))\} \quad (19)$$

$$\pi^{(\wedge)}(x|y) = \begin{cases} 1 & \text{se } \pi^{(\wedge)}(y) = \pi^{(\wedge)}(x, y) \\ \pi^{(\wedge)}(x, y) & \text{se } \pi^{(\wedge)}(y) > \pi^{(\wedge)}(x, y) \end{cases} \quad (20a)$$

$$\pi^{(\cdot)}(x|y) = \frac{\pi^{(\cdot)}(x, y)}{\pi^{(\cdot)}(y)} \quad (20b)$$

3. Exemplo

Nesta seção, as equações da metodologia (seção 2) serão aplicadas a um conjunto de dados de exemplo, com o objetivo de ilustrar a metodologia proposta. Obtém-se, para um conjunto de dados de exemplo, as funções necessárias para a aplicação do método: a distribuição priori $\pi(x)$ (seção 3.1.1) e a função de verossimilhança $\pi(y|x)$ (seção 3.1.2). A distribuição posteriori é obtida na seção 3.2, para a t-norma do mínimo (eq. 17) e produto (eq. 16).

3.1. Obtendo a priori e a verossimilhança a partir do registro histórico

A seção 2 dedicou-se a motivar a natureza do processo de inferência abordado, e a obter expressões analíticas em função da t-norma T para a distribuição

posteriori. Falta agora aplicar estas expressões para a posteriori comportam-se à algum registro histórico de exemplo e, a partir disso, obter a distribuição priori $\pi(x)$ e a verossimilhança $\pi(y|x)$. Por fim, a partir dessas duas obtém-se a posteriori.

3.1.1 Distribuição priori

Uma vez que esta a priori representa o comportamento da variável X do problema independentemente da previsão Y , a informação sobre quais valores são relevantes para a variável X (e com que pesos) pode ser fornecida, por exemplo, por um *histograma*. Escolhe-se um número fuzzy trapezoidal para modelar a função de pertinência de $\pi_X(x)$, então é preciso os parâmetros (a, u, v, b) na equação (21):

$$\text{trapez}(a, u, v, b)(x) = \begin{cases} \frac{(x-a)}{u-a} & \text{se } a \leq x < u \\ 1 & \text{se } u \leq x \leq v \\ \frac{(x-b)}{v-b} & \text{se } v \leq x < b \\ 0 & \text{caso contrário} \end{cases} \quad (21)$$

Suponha que se tenha disponível a informação sobre a ocorrência dos valores da variável X , e seja possível construir um histograma utilizando estes dados. A quantidade e o tamanho das barras do histograma foram ajustadas de forma ótima pelo programa utilizado. Informações importantes também podem ser obtidas a partir de um histograma cumulativo, que mostra de que forma o total de valores estão distribuídos. O histograma usual e o cumulativo para os dados de exemplo são mostrados a seguir na figura 2:

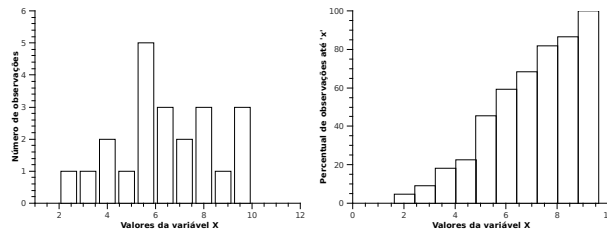


Figura 2: Histograma usual (à esquerda) e cumulativo (à direita) da variável X , usados para obter a distribuição priori $\pi_X(x)$.

Um histograma cumulativo pode servir de grande ajuda, pois informa mais facilmente o quanto cada uma das barras contribui na quantidade total de pontos. Para obtê-lo, é preciso escolher os parâmetros a, u, v e b de acordo

com a equação (21). Observando o histograma cumulativo da figura 2, nota-se que a maior quantidade de valores (o maior salto do histograma cumulativo) ocorre entre os valores $X = 5$ e $X = 8$ (40% dos valores está neste intervalo, pois a porcentagem varia de 40 a 80%).

Assim, escolhe-se que $u = 5$ e $v = 8$. De forma a englobar os outros valores, escolhe-se $a = 2$ e $b = 10$. Então, a distribuição priori, $\pi_X(x)$, é descrita pelo número fuzzy trapezoidal de parâmetros $(2, 5, 8, 10)$. Substituindo o valor na equação (21), obtém-se:

$$\pi_X(x) = \text{trapez}(2, 5, 8, 10)(x) = \begin{cases} \frac{(x-2)}{3} & \text{se } 2 < x < 5 \\ 1 & \text{se } 5 \leq x \leq 8 \\ \frac{(10-x)}{2} & \text{se } 8 < x < 10 \\ 0 & \text{caso contrário} \end{cases}, \quad (22)$$

3.1.2 Função de verossimilhança

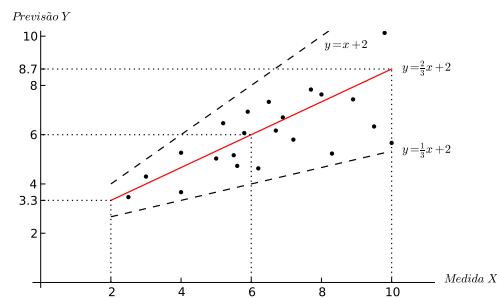


Figura 3: Gráfico $X \times Y$, utilizado para obter a função de verossimilhança $\pi_{Y|X}$

Levando em consideração que este conjunto representa a *relação* entre as medidas X e as previsões Y do problema, os parâmetros são obtidos a partir de análises de um gráfico $X \times Y$, de forma que cada par ordenado (x^*, y^*) corresponde à medida x^* e sua previsão y^* .

Para obter os parâmetros do conjunto $\pi_{Y|X}$ a partir da figura 3, é necessário o ajuste linear de mínimos quadrados dos dados e duas retas-limite (uma inferior e outra superior, em tracejado) que englobam os pontos experimentais.

As citadas “retas-limite” da figura 3 são obtidas aumentando ou diminuindo o coeficiente linear do ajuste linear até que se consiga englobar todos os pontos experimentais, e, em seguida, fornecer maior ou menor “folga” conforme a qualidade do ajuste linear (quanto melhor o ajuste, menor a folga).

As retas-limite definem na verdade o *suporte* da função de verossimilhança. Isto significa que somente os pares (x, y) que estiverem *entre* as duas retas pontilhadas têm possibilidade (valor de pertinência da verossimilhança) positivos. Pela figura, conclui-se que os pares que têm maior possibilidade (que têm possibilidade “1”) são os que pertencem ao ajuste linear do gráfico. Então modelar a verossimilhança pode resumir-se a determinar (1) as fronteiras (inferiores e superiores) de valores relevantes e (2) quais os valores de máxima possibilidade. Pode-se utilizar portanto um número fuzzy *triangular*, que é simplesmente um número fuzzy trapezoidal (eq. (21)) com $u = v$ (ou seja, com um único valor que têm pertinência “1”), de forma que:

1. As retas-limite superiores e inferiores da figura 3 forem os limites superiores e inferiores do número fuzzy triangular (respectivamente a e b) em questão;
2. O ajuste linear da figura 3 corresponda ao valor máximo do número triangular u .

Um número triangular genérico é dado pela equação:

$$\text{triang}(a, u, b)(y) = \begin{cases} \frac{(y-a)}{u-a} & \text{se } a \leq y < u \\ \frac{(y-b)}{u-b} & \text{se } u \leq y < b \\ 0 & \text{caso contrário.} \end{cases} \quad (23)$$

Dessa forma, pode-se definir a função de verossimilhança a partir de informações de um gráfico. Para este exemplo, substituindo $u = \frac{2}{3}x + 2$, $a = \frac{x}{3} + 2$ e $b = x + 2$ no número fuzzy triangular genérico (eq. 23), obtém-se²:

$$\pi_{Y|X}(y|x) = \begin{cases} 3 - 3\frac{(y-2)}{x} & \text{se } y - 2 \leq x \leq \frac{3}{2}(y - 2) \\ 3\frac{(y-2)}{x} - 1 & \text{se } \frac{3}{2}(y - 2) \leq x \leq 3(y - 2) \\ 0 & \text{caso contrário} \end{cases} \quad (24)$$

3.2. Posteriori para caso particular de previsão

Suponha que o especialista informe sobre a variável X a previsão $Y = 4$. Deseja-se obter a distribuição posteriori de X dada essa informação, para a

²Como nesta aplicação recebe-se sempre um valor de y e se trabalha com uma função de x , na eq. (23) isola-se os intervalos da expressão em termos da variável x .

distribuição priori e a função de verossimilhança obtidas no caso particular de registro histórico (eqs. (22) e (24), obtidas na seção 3.1).

A função de verossimilhança $\pi_{Y|X}(y|x)$ (24), para a previsão $y = 4$ fica:

$$\pi_{Y|X}(4|x) = \begin{cases} 3 - \frac{6}{x} & \text{se } 2 \leq x \leq 3 \\ \frac{6}{x} - 1 & \text{se } 3 < x \leq 6 \\ 0 & \text{caso contrário} \end{cases} . \quad (25)$$

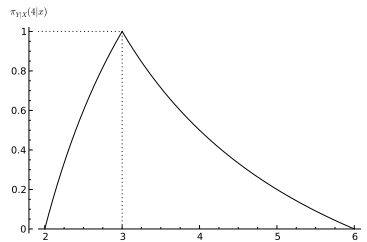


Figura 4: Gráfico da função de verossimilhança $\pi_{Y|X}(y|x)$, para o caso particular de previsão (eq. 25).

A função de verossimilhança, eq. (25), pode ser vista como: dada a previsão $Y = 4$, os valores que se espera para a variável X estão em torno de $X = 3$ (que é o valor com maior grau de pertinência). A distribuição dos valores em torno de $X = 3$ não é simétrica, pois à esquerda tem-se um intervalo de valores com pertinência maiores que zero (entre 2 e 3) menor do que à direita (entre 3 e 6). Nota-se também que o crescimento (entre 2 e 3) e o decréscimo (entre 3 e 6) ocorrem de forma não linear.

Conforme a seção 1.1, a distribuição posteriori pode ser vista como um “acordo” entre a informação da função de verossimilhança e a da distribuição priori. Para a previsão $Y = 4$, a verossimilhança de Y (eq.25) informa que o valor de X esperado é em torno de 3, com uma maior tolerância para valores maiores do que para menores que 3. Mas qual a relevância destes valores com relação à variável X , ou seja: qual sua relevância com relação à priori? Por exemplo, de acordo com a verossimilhança, o valor mais relevante é $X = 3$. Só que este valor não é tão relevante (no sentido de comum, frequente) para a variável X , pois $\pi_X(3) = 1/3$ (eq. 22). Então, pode-se dizer que dificilmente o valor $X = 3$ seria o mais relevante da distribuição posteriori $\pi_{X|Y}(x|4)$, pois o citado “acordo” que origina a posteriori deve levar em consideração a relativa baixa relevância deste valor para a distribuição priori.

A seguir, são realizados os cálculos da distribuição posteriori para os dois casos de t-norma.

3.2.1 Posteriori em $Y=4$, t-norma=produto (\cdot)

Definições da distribuição priori e da função de verossimilhança

De acordo com eq. (22) e (25):

$$\pi_X(x) = \begin{cases} \frac{(x-2)}{3} & \text{se } 2 < x < 5 \\ 1 & \text{se } 5 \leq x \leq 8 \\ \frac{(10-x)}{2} & \text{se } 8 < x < 10 \\ 0 & \text{caso contrário} \end{cases} \quad \pi_{Y|X}(4|x) = \begin{cases} 3 - \frac{6}{x} & \text{se } 2 \leq x \leq 3 \\ \frac{6}{x} - 1 & \text{se } 3 < x \leq 6 \\ 0 & \text{caso contrário} \end{cases}$$

i. Obtendo a distribuição conjunta

Partindo da fórmula (18):

$$\pi_{X,Y}(x,4) = \pi_X(x) \cdot \pi_{Y|X}(4|x) = \begin{cases} \frac{x^2-4x+4}{x} & \text{se } 2 \leq x \leq 3 \\ \frac{-x^2+8x-12}{3x} & \text{se } 3 \leq x \leq 5 \\ \frac{6}{x} - 1 & \text{se } 5 \leq x \leq 6 \end{cases} \quad (26)$$

ii. Obtendo a marginal de Y

Substituindo (26) na fórmula (19):

$$\begin{aligned} \pi_Y(4) &= \sup_x \{ \pi_{X,Y}(x,4) \}, \\ \pi_Y(4) &= \max \left(\sup_{2 \leq x \leq 3} \left\{ \frac{x^2 - 4x + 4}{x} \right\}, \sup_{3 \leq x \leq 5} \left\{ \frac{-x + 8x - 12}{3x} \right\}, \sup_{5 \leq x \leq 6} \left\{ \frac{6}{x} - 1 \right\} \right) \\ \pi_Y(4) &= \max \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{5} \right) = \frac{1}{3} \end{aligned} \quad (27)$$

iii. Obtendo a distribuição posteriori

Dividindo a eq. (26) por (27) (de acordo com eq. (20b)), obtém-se a distribuição posteriori $\pi_{X|Y}^{(\cdot)}(x|4)$:

$$\pi_{X|Y}^{(\cdot)}(x|4) = \begin{cases} \frac{3(x^2-4x+4)}{x} & \text{se } 2 \leq x < 3 \\ \frac{-x^2+8x-12}{x} & \text{se } 3 \leq x < 5 \\ \frac{18}{x} - 3 & \text{se } 5 \leq x < 6 \end{cases} \quad (28)$$

O gráfico da posteriori de X no caso fuzzy, para o caso da t-norma do produto (28), é mostrada na figura 6(a).

3.2.2 Posteriores em $Y=4$, t-norma=mínimo (\wedge)

As funções utilizadas são as mesmas dos cálculos para a t-norma do produto, ou seja, eq. 22 para a priori e 25 para a verossimilhança.

i. Obtendo a distribuição conjunta

A distribuição conjunta para a t-norma do mínimo em $Y = 4$, $\pi_{X,Y}(x, 4)$, exige o cálculo de mínimo entre as duas funções. Esta é uma tarefa que aqui será feita graficamente (Fig. 5), mas pode também ser feita de forma analítica.

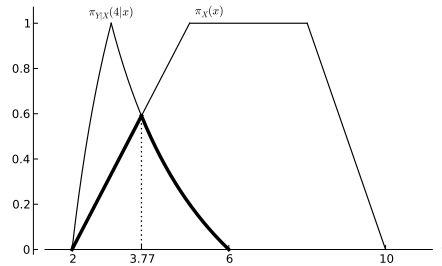


Figura 5: Figura que possibilita obter graficamente a distribuição conjunta de possibilidades para a t-norma do mínimo.

A figura 5 acima destaca o mínimo entre as duas funções, $\pi_{Y|X}(4|x)$ entre 2 e 3.77 e $\pi_X(x)$ entre 3.77 e 6, ou seja:

$$\pi_{X,Y}(x, 4) = \begin{cases} \frac{x-2}{3} & \text{se } 2 \leq x \leq 3.77 \\ \frac{6}{x} - 1 & \text{se } 3.77 \leq x \leq 6 \\ 0 & \text{caso contrário} \end{cases} \quad (29)$$

ii. Obtendo a marginal de Y

Substituindo (29) na fórmula (19):

$$\begin{aligned} \pi_Y(4) &= \sup_x \{ \pi_{X,Y}(x, 4) \} = \max \left(\sup_{2 \leq x \leq 3.77} \left\{ \frac{x-2}{3} \right\}, \sup_{3.77 \leq x \leq 6} \left\{ \frac{6}{x} - 1 \right\} \right) \\ \pi_Y(4) &= \max \left(\frac{1.77}{3}, \frac{6}{3.77} - 1 \right) = \frac{6}{3.77} - 1 \end{aligned} \quad (30)$$

iii. Obtendo a distribuição posteriori

Utilizando as eqs. (29) e (30), nota-se que $\pi_{X,Y}(3.77, 4) = \frac{6}{3.77} - 1 = \pi_Y(4)$. Então, por (20a), a distribuição posteriori para a t-norma do mínimo é:

$$\pi_{X|Y}^{(\wedge)}(x|4) = \begin{cases} \frac{x-2}{3} & \text{se } 2 \leq x < 3.77 \\ 1 & \text{se } x = 3.77 \\ \frac{6}{x} - 1 & \text{se } 3.77 < x \leq 6 \\ 0 & \text{caso contrário} \end{cases} \quad (31)$$

O gráfico da posteriori de X para a previsão $Y = 4$ para a t-norma do mínimo, $\pi_{X|Y}^{(\wedge)}$ (eq. 31), é mostrado na figura 6(b).

3.2.3 Comentários sobre o exemplo

Os gráficos das distribuições posteriori para as t-normas do produto e do mínimo, respectivamente, calculadas nas seções 3.2.1 (eq. 28) e 3.2.2 (eq. 31), utilizando distribuição priori (22) e função de verossimilhança (25) obtidas a partir de um registro histórico na seção 3.1, são mostrados abaixo na figura 6.

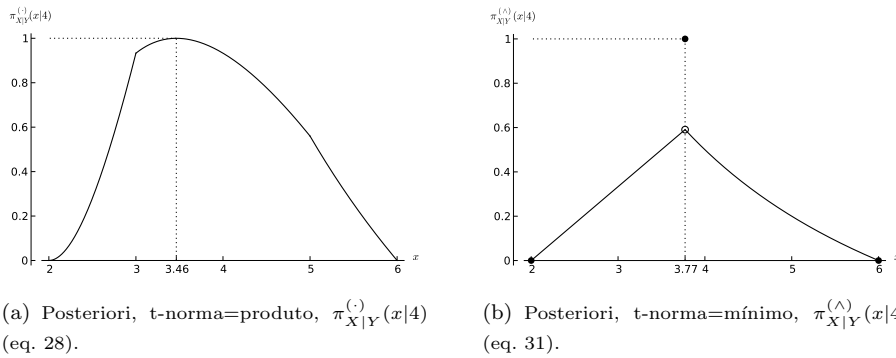


Figura 6: Distribuição posteriori para a previsão $Y = 4$, $\pi_{X|Y}(x|4)$, obtida a partir da priori e da função de verossimilhança desta seção (respectivamente, eqs. 22 e 24), para os dois casos de t-norma.

Na seção 3.2, após algumas considerações simples, parecia correto afirmar que o valor $X = 3$ não seria o mais relevante da distribuição posteriori de X .

Tal intuição foi confirmada, ao perceber que o valor mais relevante (cuja pertinência vale “1”) de ambas as distribuições posteriori (figuras 6(a) e 6(b), acima) não é $X = 3$.

Baseando-se nos valores mais relevantes de ambos os casos da posteriori, “3.46” para a t-norma do produto e “3.77” para a do mínimo (ambos destacados em seus respectivos gráficos), conclui-se então que, de acordo com a metodologia:

“Para uma previsão $Y = 4$ do especialista, o registro histórico sugere que se espera que a variável X assumira um valor *ligeiramente menor que o da previsão.*”

4. Conclusões

Propõe-se aqui um método de revisão de previsões utilizando registro histórico, que tem a vantagem de retornar soluções analíticas que têm uma grande riqueza interpretativa (pelo fato do processo de inferência ser análogo ao bayesiano). Este artigo tem o objetivo de apresentar o assunto do ponto de vista teórico e prático e motivar sua exploração e aplicação, visto que os cálculos realizados no exemplo podem abranger uma grande gama de aplicações.

Referências

- Bacani, F. (2012). Modelos de predição utilizando lógica fuzzy: uma abordagem inspirada na inferência bayesiana. Dissertação de Mestrado, IMECC–UNICAMP, Campinas/SP.
- Barros, L. C. e Bassanezi, R. C. (2010). *Tópicos de lógica fuzzy e biomatemática*, volume 5 de *Textos Didáticos*. Coleção IMECC, 2ª edição.
- Bouchon, B. (1987). Fuzzy inferences and conditional possibility distributions. *Fuzzy Sets and Systems*, 23:33–41.
- Dubois, D., Moral, S., e Prade, H. (1997). A semantics for possibility theory based on likelihoods. *J. Math. Anal. Appl.*, 205:359–380.
- Hisdal, E. (1978). Conditional possibilities independence and noninteraction. *Fuzzy Sets and Systems*, 1(4):283–297.
- Lapointe, S. e Bobée, B. (2000). Revision of possibility distributions: A bayesian inference pattern. *Fuzzy Sets and Systems*, 116(2):119–140.

- Montgomery, D. C. (1976). *Forecasting and time series analysis*. St. Louis: McGraw-Hill, New York.
- Nguyen, H. T. (1978). On conditional possibility distributions. *Fuzzy Sets and Systems*, 1(4):299–307.
- Pedrycz, W. (1982). Fuzzy relational equations with triangular norms and their resolution. *Busefal*, 11:24–32.
- Zadeh, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1(1):3–28.